Constrained Portfolio Optimization with Estimation Risk: Theory and ESG Applications

Zhenzhen Huang, Pengyu Wei, and Chengguo Weng*

Abstract

This paper studies portfolio combination for mitigating estimation risk in the first two moments of asset returns under a constrained framework. We develop a combined three-fund portfolio and derive optimal combination coefficients by maximizing expected out-of-sample mean-variance utility under expected equality or inequality constraints. The proposed sophisticated estimators allow the portfolio to attain the target factor exposure pointwise. Using environmental, social, and governance (ESG) investing as an illustration in simulation and empirical studies, we demonstrate that our combined portfolios consistently outperform standard benchmarks by delivering higher expected out-of-sample utility while reliably anchoring to the target factor exposure.

---

*Huang (corresponding author), huang.5890@osu.edu, The Ohio State University Department of Mathematics; Wei (deceased), pengyu.wei@ntu.edu.sg, Nanyang Technological University Nanyang Business School; and Weng, chengguo.weng@uwaterloo.ca, University of Waterloo Department of Statistics and Actuarial Science. We thank Thierry Foucault (the editor) and an anonymous referee for helpful comments and suggestions. Weng acknowledges financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC; grant no. RGPIN-2023-03335).

I. Introduction

Mean-variance (MV) analysis provides the theoretical foundation for modern portfolio construction, and real-world portfolio management is rarely unconstrained. Institutional mandates, regulatory requirements, and client guidelines often impose a variety of constraints, including those related to factor characteristics such as environmental, social, and governance (ESG) scores, valuation ratios, liquidity measures, or momentum exposures. Such constraints play an important role in institutional investing, especially in sustainable finance, where investors are often required to control or target the ESG profile of the portfolio. Recently, Lo and Zhang (2025) demonstrate that constraints may contain information correlated with asset returns and can therefore affect portfolio performance.

In general, the MV optimization framework assumes that the mean vector and covariance matrix of asset returns are known. In practice, however, these parameters are rarely available and have to be estimated. The standard plug-in approach first estimates these parameters from historical data and then substitutes the estimates into the optimal portfolio weights as if they were the true parameters. As a result, the plug-in portfolio may deviate substantially from the truly optimal portfolio and exhibit significant deterioration in out-of-sample (OOS) performance. This phenomenon, commonly referred to as estimation risk, has long been recognized as a central challenge in portfolio construction.

Earlier studies have devoted considerable effort to mitigating estimation errors and enhancing OOS performance in portfolio management, including works by Kalymon (1971), Klein and Bawa (1976), and Klein and Bawa (1977), among others. One important line of research develops improved estimators for unknown model parameters (Jorion (1986), (1991),

Ledoit and Wolf (2004), (2017)), while another explores additional constraints to improve OOS performance, such as no-short-sale (Jagannathan and Ma (2003)) or norm constraints (DeMiguel, Garlappi, Nogales, and Uppal (2009)). More recently, a growing literature addresses estimation risk by designing rules that maximize expected OOS utility rather than in-sample optimality, often through weighted combinations of existing portfolio rules (e.g., Kan and Zhou (2007), Tu and Zhou (2011), DeMiguel, Martín-Utrera, and Nogales (2013), (2015), Kircher and Rösch (2021), Bodnar, Okhrin, and Parolya (2023), Kan, Wang, and Zhou (2022), Lassance, Vanderveken, and Vrins (2024b), Lassance, Martín-Utrera, and Simaan (2024a), Huang, Wei, and Weng (2024), Kan and Lassance (2025)).

While most of this literature focuses on combinations of two constituent portfolios in unconstrained settings, we advance this framework by designing combination strategies that mitigate estimation risk within a constrained expected OOS MV optimization framework. Given that constraints reduce the effective degrees of freedom, mitigating estimation risk in the constrained framework generally requires combining more constituent portfolios than in the unconstrained setting. Meanwhile, because the number of combination coefficients increases with the number of constraints, there is a trade-off between the flexibility gained from including additional constituent portfolios and the estimation error associated with estimating more combination coefficients. To balance this trade-off and preserve analytical tractability, we restrict our focus to two constraints: the full-investment constraint, which requires the portfolio weights to sum to one, and a factor exposure constraint, which requires the portfolio's exposure to a factor vector, such as ESG scores or other firm-level attributes, to meet a prespecified target. Factor characteristics are treated as given at the portfolio formation stage, since they are typically observed when the investment decision is made.

We begin by analyzing the population-optimal portfolio under the MV optimization framework with the aforementioned constraints, and show that it separates into three constituent portfolios: a weight-constrained portfolio, a factor-constrained portfolio, and a zero-weight-factor portfolio. Building on the three-fund separation, we propose a combined three-fund rule whose constituent portfolios are derived from the components of the plug-in version of the population-optimal portfolio. The structure of these three constituent portfolios is designed to maintain analytical tractability while satisfying the full-investment constraint automatically and leaving two combination coefficients to be determined.

Because the combined portfolio depends on the realized sample, imposing a pointwise factor constraint, which requires the portfolio factor exposure to equal the target in every realization, generally makes constant combination coefficients infeasible. We therefore impose the factor constraint in expectation and derive the optimal combination coefficients by maximizing expected OOS MV utility under this constraint. The resulting coefficients depend on four key parameters, and we study two sets of estimators for them: sample estimators and sophisticated estimators. Under the sample estimators, the variance of the portfolio factor exposure is invariant to the target level, which means that varying the target shifts only the expected exposure without affecting its variability across estimation windows. Under the sophisticated estimators, the implementable combined portfolio interestingly satisfies the pointwise factor constraint, even though the optimization is formulated in expectation.

Compared with the combined two-fund portfolio of Kan et al. (2022), our combined three-fund portfolio introduces an additional constituent portfolio associated with the factor vector, so that the two-fund portfolio is nested as a special case of our framework. Whereas Kan et al. (2022) maximize the expected OOS MV utility under the full-investment constraint alone,

we maximize it subject to an additional expected factor constraint. Although the combined three-fund portfolio benefits from the additional constituent portfolio, the additional factor constraint offsets part of the gain. Theoretically, we characterize the target factor level at which the utility gain over the two-fund rule is the largest, and derive the range of target factor levels over which our framework attains a higher expected OOS utility.

To accommodate investors with varying preferences over factor exposure, we further consider the expected OOS MV utility maximization framework by imposing the expected factor constraint as an inequality, which allows the expected portfolio factor exposure to exceed a prespecified target. We derive the optimal combination coefficients analytically and identify a threshold for the target factor exposure. Below the threshold, the inequality constraint is inactive and the expected OOS utility remains at its unconstrained level. Once the target exceeds the threshold, the constraint binds, and both the optimal coefficients and the utility coincide with those under the equality constraint.

Furthermore, we provide two criteria for choosing the target factor level in an equality constraint: the first requires the constrained portfolio to retain a prespecified fraction of the maximum utility attainable by the unconstrained combined portfolio, and the second allows a prespecified absolute utility concession. Either criterion yields an explicit range of admissible targets centered at the threshold. An investor who favors a high factor exposure, such as an ESG investor, can choose the highest target in this range.

We apply the proposed framework to ESG investing, where our extensive simulation and empirical studies provide evidence on the effectiveness of the combined portfolios in mitigating estimation risk while accommodating the objective of meeting portfolio ESG score targets. Overall, the proposed combination approach offers ESG investors a practical way to account for

estimation risk in constrained portfolio construction, improving OOS performance while ensuring the attainment of target portfolio ESG scores.

Specifically, our numerical studies yield the following empirical findings for constrained portfolio management. First, the proposed combined portfolios consistently outperform the plug-in portfolios in terms of expected OOS MV utility, often by a substantial margin. Second, while the inequality-constrained formulation provides conceptual flexibility by permitting expected factor exposures to exceed the prespecified target, its practical implementation introduces additional estimation error in determining whether the constraint binds. Consequently, the equality-constrained portfolios generally achieve higher expected OOS utility, even at target levels where the constraint is theoretically nonbinding. The inequality formulation only regains its theoretical utility advantage when the risk aversion coefficient in the MV formulation is extremely high, or, as demonstrated in our empirical study, within smaller investment universes where overall estimation risk remains less severe. Finally, compared to traditional ESG-unaware strategies, which yield uncontrolled and volatile realized ESG scores across varying asset universe sizes, our proposed framework successfully anchors portfolios to factor target levels while delivering higher OOS performance in the large majority of cases when the two are compared at comparable ESG levels.

Having outlined our proposed methodology and its empirical efficacy, it is important to explicitly position our contributions relative to recent developments in constrained MV optimization, most notably Chavez-Bedoya and Castaneda (2024) and Wu, Lu, and Yang (2026). While these recent studies also address estimation risk under constraints, their approaches differ fundamentally from ours. Specifically, while Chavez-Bedoya and Castaneda (2024) extend the expected OOS utility framework to general equality constraints, their method of scaling

theoretical components by constants inherently violates the original constraints under sample estimates. Moreover, their work lacks simulation or empirical validation, and omits guidance for target level selection. By contrast, our framework accommodates both equality and inequality constraints, ensures that implementable strategies satisfy the target factor exposure in every realization under sophisticated estimators, and provides explicit criteria for choosing the target factor level.

Furthermore, while Wu et al. (2026) analyze a high-dimensional regularized plug-in portfolio under asymptotic settings by treating factor scores as random variables, our approach relies on exact finite-sample distributions to construct a combined portfolio rule that treats the factor as given. Fundamentally, their strategies remain plug-in portfolios by formulation, whereas our combined portfolios are optimized by maximizing expected OOS MV utility. Moreover, while the ESG score constraint is the only one in their formulation, our proposed combined portfolio additionally incorporates the full-investment constraint.

The remainder of the paper is organized as follows. Section II develops the constrained MV framework and derives the three-fund representation. Section III introduces the combined portfolio and studies its expected OOS properties. Section IV derives the optimal combination coefficients under expected factor constraints and develops implementable portfolio rules. Section V presents the simulation and empirical results for the ESG application. Finally, Section VI concludes the paper. All proofs are provided in Appendix A, and additional numerical results are reported in Appendix B.

II. Constrained Mean-Variance Framework

Consider an investor allocating wealth among $N \geq 2$ risky assets over a single period to maximize classical MV utility (Markowitz (1952)). Let $\boldsymbol{r} = [r_1, \ldots, r_N]^\top$ denote the asset return vector, $\boldsymbol{s} = [s_1, \ldots, s_N]^\top$ the associated firm characteristic vector (e.g., ESG score, P/E ratio, momentum), and $\boldsymbol{w} = [w_1, \ldots, w_N]^\top$ the portfolio weights.

From the perspective of Lo and Zhang (2025), portfolio constraints may contribute either positively or negatively to portfolio return and utility when the characteristics underlying the constraints are correlated with asset returns. Accordingly, constraints can embed economically meaningful information, and recognizing this information has important implications for performance analysis, such as understanding how investment strategies like ESG-based allocation affect portfolio efficiency. In practice, investors observe the characteristic vector $\boldsymbol{s}$ at the time of portfolio construction, whereas future asset returns are unknown. We therefore consider the portfolio optimization problem that treats $\boldsymbol{s}$ as given, and assume that the asset return vector $\boldsymbol{r}$ given the characteristic vector follows a multivariate normal distribution with a mean vector $\boldsymbol{\mu} \in \mathbb{R}^N$ and a covariance matrix $\boldsymbol{\Sigma} \in \mathbb{R}^{N \times N}$ of full rank. Assume that $\boldsymbol{\mu}$ is not collinear with $\mathbf{1}$, where $\mathbf{1}$ is an N-dimensional vector of ones, and that $\boldsymbol{\Sigma}$ is a symmetric positive definite and thus non-singular matrix.

We start by examining the problem formulation under no estimation risk. The first constraint we consider is the full-investment constraint $\boldsymbol{w}^\top \mathbf{1} = 1$. The second is a factor exposure constraint associated with the characteristic vector of the form $\boldsymbol{w}^\top \boldsymbol{s} = \bar{s}$, where $\bar{s}$ is the target level of factor exposure. Accordingly, the MV utility maximization problem subject to both the

full-investment and factor constraints is formulated as follows:[1]

$$(1) \qquad \begin{aligned} \max_{\boldsymbol{w}} \quad & U(\boldsymbol{w}) := \boldsymbol{w}^\top \boldsymbol{\mu} - \frac{\gamma}{2} \boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w}, \\ \text{s.t.} \quad & \boldsymbol{w}^\top \boldsymbol{s} = \bar{s}, \quad \boldsymbol{w}^\top \mathbf{1} = 1, \end{aligned}$$

where $\gamma > 0$ is a risk aversion coefficient.

The problem in equation (1) can be seen as a special case of the general MV problem with multiple linear equality constraints recently studied by Chavez-Bedoya and Castaneda (2024). The theoretical solution can be recovered from their general results through specific matrix operations, and we report the resulting solution and its decomposition in Lemma 1 below for completeness. While the analysis of the theoretical solution to the problem in equation (1) is not the ultimate objective of the present paper, establishing this explicit three-fund separation tailored to our specific two-constraint setup is a necessary precursor for us to design desirable combined portfolios under estimation risk.

While we focus on the equality constraint here, we extend our analysis to include an inequality factor constraint in a later section when accounting for estimation risk. Additionally, to facilitate our exposition throughout the paper, we define the matrix

$$\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}(\boldsymbol{s}\mathbf{1}^\top - \mathbf{1}\boldsymbol{s}^\top)\boldsymbol{\Sigma}^{-1},$$

which, along with its sample counterpart, is used extensively in our derivations.

[1] Given that we focus on a single-period optimization, we do not consider modeling the investor's preferences as being state non-separable, time non-separable, or both.

Lemma 1. *Under the assumption that* $\mathbf{1}^\top \boldsymbol{\Omega} \boldsymbol{s} \neq 0$*, the optimal portfolio for the problem in equation (1) is*

$$\boldsymbol{w}^*_{\text{mv}} = \boldsymbol{w}_a + \bar{s}\boldsymbol{w}_e + \frac{1}{\gamma}\boldsymbol{w}_o, \tag{2}$$

where the three constituent portfolios are defined as follows:

$$\boldsymbol{w}_a = \frac{\boldsymbol{\Omega} \boldsymbol{s}}{\mathbf{1}^\top \boldsymbol{\Omega} \boldsymbol{s}}, \quad \boldsymbol{w}_e = -\frac{\boldsymbol{\Omega} \mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Omega} \boldsymbol{s}}, \text{ and } \boldsymbol{w}_o = \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - \frac{(\mathbf{1}^\top \boldsymbol{\Omega} \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}\boldsymbol{s} - (\boldsymbol{s}^\top \boldsymbol{\Omega} \boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Omega} \boldsymbol{s}}. \tag{3}$$

Furthermore, the three constituent portfolios satisfy the following properties:

1. $\boldsymbol{w}_a$ *solves* $\min_{\boldsymbol{w}} \boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w}$ *subject to* $\boldsymbol{w}^\top \boldsymbol{s} = 0$ *and* $\boldsymbol{w}^\top \mathbf{1} = 1$*;*
2. $\boldsymbol{w}_e$ *solves* $\min_{\boldsymbol{w}} \boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w}$ *subject to* $\boldsymbol{w}^\top \boldsymbol{s} = 1$ *and* $\boldsymbol{w}^\top \mathbf{1} = 0$*;*
3. $\boldsymbol{w}_o$ *satisfies* $\boldsymbol{w}_o^\top \boldsymbol{s} = 0$ *and* $\boldsymbol{w}_o^\top \mathbf{1} = 0$*.*

According to equation (2), investors are advised to hold 100% of $\boldsymbol{w}_a$, while the exposures to $\boldsymbol{w}_e$ and $\boldsymbol{w}_o$ depend on the target factor level $\bar{s}$ and risk aversion coefficient γ. Furthermore, Lemma 1 indicates that $\boldsymbol{w}_e^\top \boldsymbol{s} = 1$ whereas $\boldsymbol{w}_a^\top \boldsymbol{s} = \boldsymbol{w}_o^\top \boldsymbol{s} = 0$, and thus, the overall factor exposure of the optimal portfolio $\boldsymbol{w}^*_{\text{mv}}$ is determined solely by the allocation to $\boldsymbol{w}_e$.

III. Estimation Risk and Combination Rules

In practice, implementing the optimal portfolio in equation (2) requires estimating the unknown parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$. Suppose we observe T periods of historical asset returns, $\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_T\}$, where $\boldsymbol{r}_t$ is the N-dimensional return vector in period t, for $t = 1, \ldots, T$, and

$T > N + 3$. Assume that $\{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_T\}$ are independent and identically distributed, and each follows a multivariate normal distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. Then, it is well known that the sample mean $\hat{\boldsymbol{\mu}}$ and sample covariance matrix $\hat{\boldsymbol{\Sigma}}$ are independent of each other, with distributions given by:

$$\hat{\boldsymbol{\mu}} = \frac{1}{T}\sum_{t=1}^{T} \boldsymbol{r}_t \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/T) \quad \text{and} \quad \hat{\boldsymbol{\Sigma}} = \frac{1}{T}\sum_{t=1}^{T} (\boldsymbol{r}_t - \hat{\boldsymbol{\mu}})(\boldsymbol{r}_t - \hat{\boldsymbol{\mu}})^\top \sim \mathcal{W}_N(T-1, \boldsymbol{\Sigma})/T,$$

where $\mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/T)$ denotes the multivariate normal distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}/T$, and $\mathcal{W}_N(T-1, \boldsymbol{\Sigma})$ represents the Wishart distribution with $T-1$ degrees of freedom and scaling matrix $\boldsymbol{\Sigma}$.

The plug-in portfolio, denoted by $\hat{\boldsymbol{w}}_{\text{mv}}$, is obtained by replacing the true parameters $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ in equation (2) with their sample counterparts $\hat{\boldsymbol{\mu}}$ and $\hat{\boldsymbol{\Sigma}}$, respectively. While readily implementable, the plug-in approach overlooks the detrimental impacts of estimation errors in the unknown parameters, which leads to substantial deterioration in OOS performance relative to the truly optimal portfolio $\boldsymbol{w}^*_{\text{mv}}$.

To mitigate the estimation risk inherent in the plug-in portfolio, we focus on a widely studied class of investment strategies known as combined portfolios. In the existing literature, the components of a combined portfolio are typically the same as those in the plug-in portfolio, with combination coefficients determined under specific optimization frameworks (e.g., Kan et al. (2022)). While a natural starting point is a direct weighted combination of the three components of the plug-in portfolio $\hat{\boldsymbol{w}}_{\text{mv}}$, the complex structure of these constituents renders the analytical evaluation of their expected OOS performance intractable. In view of this challenge, we propose

the following configuration for the combined portfolio:

$$\hat{\boldsymbol{w}}_c = \hat{\boldsymbol{w}}_g + \delta\hat{\boldsymbol{w}}_z + \beta\hat{\boldsymbol{w}}_s, \tag{4}$$

where δ and β are two constant combination coefficients, and the constituent portfolios are

$$\hat{\boldsymbol{w}}_g = \frac{\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}}{\mathbf{1}^\top\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}}, \quad \hat{\boldsymbol{w}}_z = \hat{\boldsymbol{\Upsilon}}\hat{\boldsymbol{\mu}} \quad \text{and} \quad \hat{\boldsymbol{w}}_s = \hat{\boldsymbol{\Upsilon}}\boldsymbol{s},$$

with $\hat{\boldsymbol{\Upsilon}} = \hat{\boldsymbol{\Sigma}}^{-1} - \hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}\mathbf{1}^\top\hat{\boldsymbol{\Sigma}}^{-1}/\mathbf{1}^\top\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}$.

By construction, the first two constituent portfolios ($\hat{\boldsymbol{w}}_g$ and $\hat{\boldsymbol{w}}_z$) are the plug-in global minimum variance portfolio and a plug-in zero-investment portfolio, which are exactly the components of the combined two-fund rule studied by Kan et al. (2022). The third zero-investment constituent portfolio, $\hat{\boldsymbol{w}}_s$, directed by the factor vector $\boldsymbol{s}$, is introduced to accommodate the factor constraint. Because $\hat{\boldsymbol{w}}_g^\top\mathbf{1} = 1$ while $\hat{\boldsymbol{w}}_z^\top\mathbf{1} = \hat{\boldsymbol{w}}_s^\top\mathbf{1} = 0$, the combined portfolio automatically satisfies the full-investment constraint ($\hat{\boldsymbol{w}}_c^\top\mathbf{1} = 1$) for any choice of δ and β. The factor exposure constraint is deliberately left unrestricted at this stage. It will be formally enforced in Section IV when optimizing the combination coefficients. Furthermore, this configuration can recover the plug-in portfolio $\hat{\boldsymbol{w}}_{\text{mv}}$ through a specific pair of coefficients δ and β.

It is worth contrasting our design in equation (4) with the framework of Chavez-Bedoya and Castaneda (2024). Their approach separates the theoretically optimal solution into two parts and scales their plug-in versions directly, which inherently violates portfolio constraints. In contrast, our constituents $\hat{\boldsymbol{w}}_g$, $\hat{\boldsymbol{w}}_z$, and $\hat{\boldsymbol{w}}_s$ are deliberately engineered to circumvent this issue. As shown above, this construction strictly preserves the full-investment constraint automatically,

leaving the combination coefficients δ and β free to be optimized subject to the expected factor exposure target.

To facilitate the evaluation of the expected OOS MV utility for the combined portfolio $\hat{\boldsymbol{w}}_c$ in subsequent sections, we derive the analytical expressions for its expected return, expected factor exposure, and expected variance in Proposition 1 below.

Proposition 1. *For $T > N + 3$, we have*

$$\begin{aligned} \mathbb{E}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{\mu}] &= \mu_g + k_1 \delta \theta_1^2 + k_1 \beta \theta_2^2, \\ \mathbb{E}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{s}] &= \mu_s + k_1 \delta \theta_2^2 + k_1 \beta \theta_3^2, \\ \mathbb{E}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{\Sigma} \hat{\boldsymbol{w}}_c] &= k_1 k_3 [\sigma_g^2 + T k_2 (\delta^2 \theta_1^2 + \beta^2 \theta_3^2 + 2\delta\beta\theta_2^2) + k_2 (N-1)\delta^2], \end{aligned} \tag{5}$$

where $k_1 = T/(T-N-1)$, $k_2 = T/[(T-N)(T-N-3)]$, $k_3 = (T-2)/T$, *and*

$$\begin{aligned} \theta_1^2 &= \boldsymbol{\mu}^\top \boldsymbol{\Upsilon} \boldsymbol{\mu}, \qquad \theta_2^2 = \boldsymbol{\mu}^\top \boldsymbol{\Upsilon} \boldsymbol{s}, \qquad \theta_3^2 = \boldsymbol{s}^\top \boldsymbol{\Upsilon} \boldsymbol{s}, \\ \sigma_g^2 &= \frac{1}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}}, \qquad \mu_g = \frac{\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}}, \qquad \mu_s = \frac{\boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}}, \end{aligned} \tag{6}$$

with $\boldsymbol{\Upsilon} = \boldsymbol{\Sigma}^{-1} - \boldsymbol{\Sigma}^{-1} \mathbf{1} \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} / \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}$.

Additionally, we compute the variance of the factor exposure, $\hat{\boldsymbol{w}}_c^\top \boldsymbol{s}$, in Proposition 2, which will be utilized in our subsequent analysis.

Proposition 2. *The variance of $\hat{\boldsymbol{w}}_c^\top \boldsymbol{s}$ is given by*

$$\mathrm{Var}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{s}] = k_1 \theta_3^2 \left[\frac{\sigma_g^2}{T} + \delta^2 k_2 k_3 + k_2 (\delta^2 \theta_1^2 + \beta^2 \theta_3^2 + 2\delta\beta\theta_2^2) \right] + (T k_2 - k_1^2) \left(\delta \theta_2^2 + \beta \theta_3^2 \right)^2. \tag{7}$$

IV. Optimization with Expected Factor Constraints

In this section, we determine the optimal combination coefficients for the combined portfolio $\hat{\boldsymbol{w}}_c$ in equation (4) by adopting the expected OOS MV utility optimization framework of Kan and Zhou (2007). We first derive the optimal combined portfolio under an equality factor constraint, formulate its implementable version, and then extend our analysis to an inequality-constrained framework to accommodate more flexible factor exposure preferences. Finally, we provide explicit criteria to guide the selection of the target factor level.

A. Optimal Combined Portfolio

For the factor exposure target, a natural starting point is to impose a pointwise equality constraint, as in the problem under no estimation risk discussed in Section II, i.e., $\hat{\boldsymbol{w}}_c^\top \boldsymbol{s} = \bar{s}$. However, because the constituent portfolios in $\hat{\boldsymbol{w}}_c$ are estimated from historical data, they are inherently random. Consequently, no constant coefficients δ and β can satisfy this strict equality almost surely across all sample realizations, which renders the pointwise constraint infeasible. To circumvent this issue, we relax the pointwise requirement and instead enforce the factor constraint in expectation. Assuming investors mandate that their portfolio's average factor exposure exactly matches a prespecified target $\bar{s}$, we impose the strict equality constraint $\mathbb{E}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{s}] = \bar{s}$. This leads to the following expected OOS MV utility maximization problem:

$$
\begin{aligned}
(8) \qquad \max_{\delta,\beta} \quad & \mathbb{E}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{\mu}] - \frac{\gamma}{2}\mathbb{E}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{\Sigma} \hat{\boldsymbol{w}}_c], \\
\text{s.t.} \quad & \mathbb{E}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{s}] = \bar{s}.
\end{aligned}
$$

The optimal combination coefficients for the problem in equation (8) can be obtained using the method of Lagrange multipliers. The result is summarized in Proposition 3, where we use the subscript 'E' to denote the optimal combination for the equality-constrained formulation in equation (8), distinguishing it from the inequality-constrained case that will be studied in a later section.

Proposition 3. *The optimal combination coefficients δ_E^* and β_E^* that solve the problem in equation (8) are given by*

$$\delta_E^* = \frac{\theta_1^2\theta_3^2 - \theta_2^4}{\gamma k_2 k_3 D} \quad \textit{and} \quad \beta_E^* = -\frac{k_1\theta_2^2\delta_E^* + (\mu_s - \bar{s})}{k_1\theta_3^2}, \tag{9}$$

where $D = T(\theta_1^2\theta_3^2 - \theta_2^4) + \theta_3^2(N-1)$. The corresponding optimal combined portfolio is

$$\hat{\boldsymbol{w}}_{c\text{-}E}^* = \hat{\boldsymbol{w}}_g + \delta_E^*\hat{\boldsymbol{w}}_z + \beta_E^*\hat{\boldsymbol{w}}_s. \tag{10}$$

Our constrained three-fund formulation introduces a structural trade-off relative to the two-fund framework of Kan et al. (2022). On the one hand, introducing the third constituent $\hat{\boldsymbol{w}}_s$ expands the feasible set, which inherently enhances the maximum achievable utility. On the other hand, imposing the expected factor constraint restricts this set, penalizing the utility. Proposition 4 evaluates this net effect by deriving the exact difference in expected OOS MV utility between our constrained three-fund portfolio and the unconstrained two-fund benchmark.

Proposition 4. *Let $\hat{\boldsymbol{w}}_k^*$ denote the optimal two-fund portfolio for the unconstrained utility maximization problem in Kan et al. (2022). Then the difference in expected OOS MV utility*

between $\hat{\boldsymbol{w}}^*_{c\text{-}E}$ *defined in equation* (10) *and* $\hat{\boldsymbol{w}}^*_k$ *is*

$$\Delta_U := \mathbb{E}[U(\hat{\boldsymbol{w}}^*_{c\text{-}E})] - \mathbb{E}[U(\hat{\boldsymbol{w}}^*_k)] = -\frac{\gamma k_2 k_3 T}{2k_1\theta_3^2}\left(\bar{s} - \bar{s}^*\right)^2 + \frac{k_1\theta_2^4(N-1)^2}{2\gamma k_2 k_3 TAD}, \tag{11}$$

where $A = T\theta_1^2 + N - 1$ *and*

$$\bar{s}^* = \mu_s + \frac{k_1\theta_2^2}{\gamma k_2 k_3 T}. \tag{12}$$

The utility difference Δ_U in equation (11) is a concave quadratic function of the target factor level $\bar{s}$, and is maximized at $\bar{s}^*$. Consequently, $\bar{s}^*$ represents the expected factor exposure of the optimal three-fund portfolio when the factor constraint in equation (8) is absent. Any target deviation ($\bar{s} \neq \bar{s}^*$) penalizes the utility difference by an amount proportional to $(\bar{s} - \bar{s}^*)^2$. Meanwhile, given $A > 0$ and $D > 0$, the second term in equation (11) is non-negative and represents the maximum achievable utility gain of the three-fund portfolio over the two-fund benchmark. This gain is strictly positive whenever $\theta_2^2 = \boldsymbol{\mu}^\top \boldsymbol{\Upsilon} \boldsymbol{s} \neq 0$. This demonstrates that the structural advantage of including the third constituent $\hat{\boldsymbol{w}}_s$ is fundamentally driven by the alignment between the expected asset return vector and the factor vector.

Based on the utility difference established in Proposition 4, Corollary 1 below formalizes the conditions under which our constrained three-fund portfolio dominates the unconstrained two-fund benchmark.

Corollary 1. *We have the following assertions:*

1. *A non-empty dominance region* $\{\bar{s} : \Delta_U > 0\}$ *exists if and only if* $\theta_2^2 \neq 0$.

2. *When* $\theta_2^2 \neq 0$*, the optimal three-fund portfolio outperforms the optimal two-fund portfolio*

in Kan et al. (2022) if and only if

$$|\bar{s} - \bar{s}^*| < \frac{(N-1)k_1}{\gamma T k_2 k_3} \sqrt{\frac{\theta_2^4 \theta_3^2}{AD}}. \tag{13}$$

Corollary 1 provides important economic intuition for the dominance region. If the factor characteristic is uninformative with respect to the asset returns ($\theta_2^2 = 0$), the peak $\bar{s}^*$ reduces to μ_s, and the three-fund portfolio offers no utility gain over the two-fund benchmark ($\Delta_U \leq 0$). However, when $\theta_2^2 \neq 0$, the dominance region forms a symmetric interval around $\bar{s}^*$. Economically, a wider interval implies that the factor component is highly informative and allows the investor to sustain a broader range of target exposures without sacrificing utility relative to the unconstrained two-fund portfolio.

B. Implementable Combined Portfolio

In practice, the optimal combination coefficients in Proposition 3 are not directly implementable because they depend on four unknown population quantities: θ_1^2, θ_2^2, θ_3^2, and μ_s. To construct an implementable portfolio, a natural approach is to estimate these parameters by substituting the sample mean vector and sample covariance matrix into their theoretical expressions, yielding the following sample estimators:

$$\hat{\theta}_1^2 = \hat{\boldsymbol{\mu}}^\top \hat{\boldsymbol{\Upsilon}} \hat{\boldsymbol{\mu}}, \;\; \hat{\theta}_2^2 = \hat{\boldsymbol{\mu}}^\top \hat{\boldsymbol{\Upsilon}} \boldsymbol{s}, \;\; \hat{\theta}_3^2 = \boldsymbol{s}^\top \hat{\boldsymbol{\Upsilon}} \boldsymbol{s} \text{ and } \hat{\mu}_s = \frac{\boldsymbol{s}^\top \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{1}}{\mathbf{1}^\top \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{1}}. \tag{14}$$

However, these sample estimators can exhibit significant bias, especially when T is small relative to N (e.g., Kan and Zhou (2007)). We therefore also consider what we refer to as

sophisticated estimators, namely analytically bias-adjusted estimators for θ_1^2, θ_2^2 and θ_3^2 constructed using the distributional properties of their sample counterparts. Additionally, due to the difficulty in obtaining a sophisticated estimator for μ_s, we opt to keep the sample estimator for this quantity. Unlike $\hat{\theta}_1^2, \hat{\theta}_2^2$ and $\hat{\theta}_3^2$, the sample estimator for μ_s is unbiased.

The sophisticated estimator for θ_1^2 as suggested by Kan et al. (2022) is given by

$$\tilde{\theta}_1^2 = \frac{(T-N-1)\hat{\theta}_1^2 - (N-1)}{T} + \frac{2(\hat{\theta}_1^2)^{\frac{N-1}{2}}(1+\hat{\theta}_1^2)^{-\frac{T-2}{2}}}{T B_{\hat{\theta}_1^2/(1+\hat{\theta}_1^2)}((N-1)/2, (T-N+1)/2)}, \tag{15a}$$

where $B_x(a,b) = \int_0^x y^{a-1}(1-y)^{b-1}\,\mathrm{d}y$ denotes the incomplete beta function. The first term in equation (15a) is the unbiased estimator of θ_1^2, while the second term provides an additional adjustment for a small estimation window T. Note that $\tilde{\theta}_1^2$ is always positive, unlike the unbiased estimator of θ_1^2. We also derive the following sophisticated (unbiased) estimators for θ_2^2 and θ_3^2:

$$\tilde{\theta}_2^2 = \frac{T-N-1}{T}\hat{\theta}_2^2 \quad \text{and} \quad \tilde{\theta}_3^2 = \frac{T-N-1}{T}\hat{\theta}_3^2. \tag{15b}$$

A detailed derivation of these two unbiased estimators is provided in Appendix A.7.

Therefore, after obtaining estimates of the four unknown quantities θ_1^2, θ_2^2, θ_3^2, and μ_s, we substitute them into the optimal combination coefficients to obtain implementable versions of the optimal combined portfolio $\hat{\boldsymbol{w}}_{c\text{-}E}^*$ in equation (10). Let $\hat{\delta}_E$ and $\hat{\beta}_E$ denote the combination coefficients estimated with the sample estimators in equation (14), and $\tilde{\delta}_E$ and $\tilde{\beta}_E$ those estimated with the sophisticated estimators in equation (15). The corresponding implementable combined portfolios are $\hat{\boldsymbol{w}}_{c\text{-}E} = \hat{\boldsymbol{w}}_g + \hat{\delta}_E\hat{\boldsymbol{w}}_z + \hat{\beta}_E\hat{\boldsymbol{w}}_s$ and $\tilde{\boldsymbol{w}}_{c\text{-}E} = \hat{\boldsymbol{w}}_g + \tilde{\delta}_E\hat{\boldsymbol{w}}_z + \tilde{\beta}_E\hat{\boldsymbol{w}}_s$, respectively. We summarize two notable properties of these implementable portfolios in Proposition 5.

Proposition 5. *For the implementable portfolios of $\hat{\boldsymbol{w}}^*_{c\text{-}E}$, we have the following assertions:*

1. *When the sample estimators in equation (14) are employed, the variance of the realized portfolio factor exposure, $\mathrm{Var}[\hat{\boldsymbol{w}}^\top_{c\text{-}E}\boldsymbol{s}]$, is invariant to the target factor level $\bar{s}$.*

2. *When the sophisticated estimators in equation (15) are used, the portfolio factor exposure satisfies the constraint pointwise: $\tilde{\boldsymbol{w}}^\top_{c\text{-}E}\boldsymbol{s} = \bar{s}$.*

While the formal proof is detailed in Appendix A.8, the mathematical mechanics driving the results in Proposition 5 can be outlined as follows. First, for $\hat{\boldsymbol{w}}_{c\text{-}E}$ constructed using the sample estimators, the portfolio factor exposure, $\hat{\boldsymbol{w}}^\top_{c\text{-}E}\boldsymbol{s}$, is a linear function of the target factor level $\bar{s}$ with a constant slope. As a result, varying $\bar{s}$ changes the expected level of the portfolio factor exposure, but does not affect its variability across estimation windows. Second, when the sophisticated estimators are used, the resulting portfolio $\tilde{\boldsymbol{w}}_{c\text{-}E}$ eliminates this variability altogether, satisfying the pointwise constraint $\tilde{\boldsymbol{w}}^\top_{c\text{-}E}\boldsymbol{s} = \bar{s}$ exactly for every estimation window. This remarkable cancellation is driven by the structural unbiasedness of $\tilde{\theta}^2_2$ and $\tilde{\theta}^2_3$. Specifically, because one combination coefficient is derived from the other via the expected factor constraint, the bias adjustments in the sophisticated estimators cause the relevant sample terms to cancel out perfectly when evaluating $\tilde{\boldsymbol{w}}^\top_{c\text{-}E}\boldsymbol{s}$, and leave only the target level $\bar{s}$. Therefore, although the optimal combined portfolio was originally derived under the expected constraint $\mathbb{E}[(\hat{\boldsymbol{w}}^*_{c\text{-}E})^\top\boldsymbol{s}] = \bar{s}$, the implementable portfolio based on the sophisticated estimators, interestingly, satisfies the corresponding pointwise factor constraint.

C. Extension to an Inequality Factor Constraint

To accommodate investors who only require a minimum factor threshold, we relax the strict equality to an inequality constraint. This formulation strictly expands the feasible set, potentially enhancing the maximum achievable utility. Accordingly, we consider the following generalized optimization framework:

$$\begin{aligned} \max_{\delta,\beta} \quad & \mathbb{E}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{\mu}] - \frac{\gamma}{2}\mathbb{E}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{\Sigma} \hat{\boldsymbol{w}}_c], \\ \text{s.t.} \quad & \mathbb{E}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{s}] \geq \bar{s}. \end{aligned} \tag{16}$$

We solve the optimization problem in equation (16) using the Karush–Kuhn–Tucker conditions. The result is summarized in Proposition 6.

Proposition 6. *The optimal combination coefficients δ_I^* and β_I^* that solve the problem in equation (16) are given by*

$$\delta_I^* = \delta_E^* \quad \textit{and} \quad \beta_I^* = \begin{cases} \beta_E^*, & \zeta \geq 0, \\ \frac{(N-1)\theta_2^2}{\gamma k_2 k_3 T D}, & \zeta < 0, \end{cases} \tag{17}$$

where $\zeta = -[k_1\theta_2^2 + \gamma k_2 k_3 T(\mu_s - \bar{s})]/(k_1\theta_3^2)$. The corresponding optimal combined portfolio is

$$\hat{\boldsymbol{w}}_{c\text{-}I}^* = \hat{\boldsymbol{w}}_g + \delta_I^* \hat{\boldsymbol{w}}_z + \beta_I^* \hat{\boldsymbol{w}}_s. \tag{18}$$

Proposition 6 shows that the parameter ζ determines whether the inequality constraint

binds. In particular, the condition $\zeta = 0$ is equivalent to $\bar{s} = \bar{s}^*$, where $\bar{s}^*$ is the peak factor exposure defined in equation (12). Corollary 2 formalizes this structural relationship.

Corollary 2. *By the condition $\zeta = 0$, we obtain the peak factor exposure $\bar{s}^*$ defined in equation (12). Furthermore, we have the following assertions:*

1. *If $\bar{s} < \bar{s}^*$, the optimal combined portfolio $\hat{\boldsymbol{w}}_{c\text{-}I}^*$ in equation (18) satisfies $\mathbb{E}[(\hat{\boldsymbol{w}}_{c\text{-}I}^*)^\top \boldsymbol{s}] > \bar{s}$.*
2. *If $\bar{s} \geq \bar{s}^*$, the optimal combined portfolio $\hat{\boldsymbol{w}}_{c\text{-}I}^*$ in equation (18) satisfies $\mathbb{E}[(\hat{\boldsymbol{w}}_{c\text{-}I}^*)^\top \boldsymbol{s}] = \bar{s}$.*

Corollary 2 provides a clear economic interpretation of the optimal strategy. When $\bar{s} < \bar{s}^*$ (equivalently, $\zeta < 0$), the unconstrained portfolio naturally achieves a factor exposure that exceeds the investor's minimum requirement. Consequently, the constraint is inactive, and the portfolio allocation is driven entirely by the standard trade-off between expected OOS mean and variance. Conversely, when $\bar{s} \geq \bar{s}^*$ (equivalently, $\zeta \geq 0$), the investor demands a factor exposure at least as high as what the unconstrained optimum provides. The constraint becomes binding, forcing the optimal solution to the boundary of the feasible set, thereby reducing the problem exactly to the equality-constrained framework studied in Section IV.A.

D. Selection of the Target Factor Exposure

To guide the practical selection of the target factor exposure $\bar{s}$, we provide two criteria that explicitly quantify the trade-off between the chosen target and its associated utility cost. Let U_{unc} denote the maximum expected OOS MV utility attainable by the combined portfolio in equation (4) in the absence of the factor constraint. This unconstrained maximum is achieved when the constraint is inactive ($\zeta < 0$ in Proposition 6), corresponding exactly to the peak

exposure $\bar{s}^*$, and can be computed as follows:

$$U_{\text{unc}} := \max_{\delta,\beta} \mathbb{E}[U(\hat{\boldsymbol{w}}_c)] = \mu_g - \frac{\gamma}{2} k_1 k_3 \sigma_g^2 + \frac{k_1 \left[T\theta_1^2(\theta_1^2\theta_3^2 - \theta_2^4) + (N-1)\theta_2^4\right]}{2\gamma k_2 k_3 T D}.$$

Furthermore, the expected OOS MV utility under the equality factor constraint can be computed and expressed as

$$\mathbb{E}[U(\hat{\boldsymbol{w}}^*_{c\text{-}E})] = U_{\text{unc}} - \frac{\gamma T k_2 k_3}{2k_1\theta_3^2} \left(\bar{s} - \bar{s}^*\right)^2. \tag{19}$$

The second term of equation (19) is the utility cost of the target. It is zero at $\bar{s} = \bar{s}^*$ and increases with the distance from $\bar{s}^*$.

By specifying an acceptable utility threshold, the investor can determine the corresponding admissible range for the target factor exposure from equation (19). We define this threshold using two practical criteria: a relative criterion that requires the portfolio to retain a prespecified fraction α (e.g., 90%) of the unconstrained maximum U_{unc}, and an absolute criterion that permits a maximum utility concession d. Proposition 7 formalizes the exact target range under each criterion.

Proposition 7. *We provide the following two criteria to guide the choice of the target $\bar{s}$:*

1. *Suppose $U_{\text{unc}} > 0$ and let $\alpha \in [0,1]$. Then $\mathbb{E}[U(\hat{\boldsymbol{w}}^*_{c\text{-}E})] \geq \alpha\, U_{\text{unc}}$ if and only if*

$$|\bar{s} - \bar{s}^*| \leq \sqrt{\frac{2k_1\theta_3^2(1-\alpha)U_{\text{unc}}}{\gamma k_2 k_3 T}}. \tag{20}$$

Furthermore, equality holds in equation (20) if and only if $\mathbb{E}[U(\hat{\boldsymbol{w}}^*_{c\text{-}E})] = \alpha\, U_{\text{unc}}$.

2. *For any utility concession* $d \geq 0$, $\mathbb{E}[U(\hat{\boldsymbol{w}}^*_{c\text{-}E})] \geq U_{\text{unc}} - d$ *if and only if*

$$|\bar{s} - \bar{s}^*| \leq \sqrt{\frac{2k_1\theta_3^2 d}{\gamma k_2 k_3 T}}. \tag{21}$$

Furthermore, equality holds in equation (21) if and only if $\mathbb{E}[U(\hat{\boldsymbol{w}}^*_{c\text{-}E})] = U_{\text{unc}} - d$.

The intervals for $\bar{s}$ derived in Proposition 7 encompass all factor targets that satisfy the specified utility range. Because the utility function in equation (19) is symmetric around the peak exposure $\bar{s}^*$, the two endpoints of each interval yield exactly the minimum acceptable utility (either αU_{unc} or $U_{\text{unc}} - d$), while any interior target yields a utility strictly above it. A purely MV investor indifferent to the factor would simply maximize utility by selecting the interior peak $\bar{s}^*$. In contrast, an investor with directional factor preferences will gravitate toward the boundaries to maximize or minimize exposure subject to their utility budget. For instance, an investor seeking high factor exposure, such as the ESG-motivated investor in Pedersen, Fitzgibbons, and Pomorski (2021), will select the upper endpoint. Under the relative criterion with $\alpha = 0.90$ in equation (20), this upper bound represents the maximum attainable factor exposure that still preserves 90% of the unconstrained maximum utility. Conversely, an investor seeking to minimize exposure will select the lower endpoint. In the limiting case where the investor demands maximum utility without compromise ($\alpha = 1$ or $d = 0$), the interval collapses exactly to the unconstrained optimum $\bar{s}^*$.

Furthermore, Corollary 1 can be interpreted as a special case of the absolute criterion in Proposition 7. Specifically, if the maximum permissible utility concession d is set to be the constant term in equation (11), the acceptable utility threshold $U_{\text{unc}} - d$ coincides exactly with the benchmark utility $\mathbb{E}[U(\hat{\boldsymbol{w}}^*_k)]$. Consequently, enforcing $\mathbb{E}[U(\hat{\boldsymbol{w}}^*_{c\text{-}E})] \geq U_{\text{unc}} - d$ guarantees that the

proposed three-fund portfolio achieves an expected utility no lower than that of the two-fund benchmark (i.e., $\Delta_U \geq 0$). Substituting this specific d into equation (21) and excluding the two endpoints recovers the exact interval previously derived in equation (13). At the boundaries of this interval, the three-fund and two-fund portfolios deliver identical expected OOS MV utility.

V. Numerical Studies for ESG Applications

In this section, we apply the proposed constrained MV framework to ESG investing by defining the factor vector $\boldsymbol{s}$ as the assets' ESG scores. We evaluate performance using both simulated data under the multivariate normality assumption to verify theoretical properties, and real market data to demonstrate practical OOS efficacy.

A. The Setup and Portfolio Strategies

While we illustrate our methodology using ESG investing, the proposed framework accommodates any quantitative firm characteristic. Accordingly, we treat the factor scores as given and do not analyze subjective investor preferences. Furthermore, while ESG ratings often exhibit dispersion across rating agencies (see, for example, Berg, Koelbel, and Rigobon (2022), Avramov, Cheng, Lioui, and Tarelli (2022), Luo, Shan, Yan, and Yan (2023), among others), we do not address this underlying disagreement. Rather, we assume that the investor adopts a specific agency's ratings as a benchmark and uses these scores consistently throughout the portfolio construction process.

Given a portfolio $\boldsymbol{w}$, we measure its aggregate ESG profile as the weighted average $\boldsymbol{w}^\top \boldsymbol{s}$, with higher scores indicating better alignment with sustainability practices. For assets held in long

positions, their ESG scores positively contribute to the portfolio ESG score. Conversely, holding an asset in a short position acts as a bet against that company's sustainability. Consequently, we assume that the investor mandates this aggregate exposure to meet a prespecified target level $\bar{s}$, where a higher $\bar{s}$ indicates a stronger ESG preference.

We now detail the portfolio strategies evaluated in our numerical studies. While the previous theoretical analysis relies on the sample covariance matrix of asset returns, practical implementation often benefits from adopting shrinkage estimators of the covariance matrix. We consider the linear shrinkage estimator of Ledoit and Wolf (2004) here, while relegating results using the nonlinear shrinkage estimator of Ledoit and Wolf (2020) to Appendices B.2.4 and B.3.2 for the simulation and empirical studies, respectively.

Using the linear shrinkage estimator $\hat{\boldsymbol{\Sigma}}_{\text{LW}}$, where the subscript LW stands for Ledoit and Wolf (2004), we construct two sets of parameter estimators for the optimal combination coefficients. First, substituting $\hat{\boldsymbol{\Sigma}}_{\text{LW}}$ and the sample mean vector $\hat{\boldsymbol{\mu}}$ into equation (6) yields the shrinkage sample estimators:

$$\hat{\theta}^2_{1,\text{LW}} = \hat{\boldsymbol{\mu}}^\top \hat{\boldsymbol{\Upsilon}}_{\text{LW}} \hat{\boldsymbol{\mu}}, \ \hat{\theta}^2_{2,\text{LW}} = \hat{\boldsymbol{\mu}}^\top \hat{\boldsymbol{\Upsilon}}_{\text{LW}} \boldsymbol{s}, \ \hat{\theta}^2_{3,\text{LW}} = \boldsymbol{s}^\top \hat{\boldsymbol{\Upsilon}}_{\text{LW}} \boldsymbol{s} \text{ and } \hat{\mu}_{s,\text{LW}} = \frac{\boldsymbol{s}^\top \hat{\boldsymbol{\Sigma}}^{-1}_{\text{LW}} \mathbf{1}}{\mathbf{1}^\top \hat{\boldsymbol{\Sigma}}^{-1}_{\text{LW}} \mathbf{1}}, \tag{22}$$

where $\hat{\boldsymbol{\Upsilon}}_{\text{LW}} = \hat{\boldsymbol{\Sigma}}^{-1}_{\text{LW}} - \hat{\boldsymbol{\Sigma}}^{-1}_{\text{LW}} \mathbf{1}\mathbf{1}^\top \hat{\boldsymbol{\Sigma}}^{-1}_{\text{LW}} / \mathbf{1}^\top \hat{\boldsymbol{\Sigma}}^{-1}_{\text{LW}} \mathbf{1}$. Second, we replace the sample estimators in equation (15) with the above shrinkage sample estimators to obtain a set of shrinkage sophisticated estimators:

$$\tilde{\theta}^2_{1,\text{LW}}, \ \tilde{\theta}^2_{2,\text{LW}} \text{ and } \tilde{\theta}^2_{3,\text{LW}}. \tag{23}$$

We retain the shrinkage sample estimator for μ_s rather than introducing a separate sophisticated counterpart.

For benchmark portfolios, we consider the plug-in portfolios of equation (2) alongside ESG-constrained portfolios located on the MV efficient frontier (EF). Reconciling the EF portfolios with the ESG constraint represents a recognized approach to mitigating estimation risk relative to the globally optimal portfolio (Simaan (2014), Lassance (2022)). Imposing the target $\boldsymbol{w}_{\text{ef}}^{\top}\boldsymbol{s} = \bar{s}$ on the theoretical EF portfolio $\boldsymbol{w}_{\text{ef}} = \boldsymbol{w}_g + c\boldsymbol{w}_z$ directly yields $c = (\bar{s} - \mu_s)/\theta_2^2$. For our numerical studies, we adopt the implementable version:

$$\hat{\boldsymbol{w}}_{\text{ef}} = \hat{\boldsymbol{w}}_g + \hat{c}\hat{\boldsymbol{w}}_z, \tag{24}$$

where $\hat{c} = (\bar{s} - \hat{\mu}_s)/\hat{\theta}_2^2$. We deliberately employ the sample estimator $\hat{\theta}_2^2$, rather than the sophisticated $\tilde{\theta}_2^2$, to guarantee that $\hat{\boldsymbol{w}}_{\text{ef}}$ satisfies the constraint $\hat{\boldsymbol{w}}_{\text{ef}}^{\top}\boldsymbol{s} = \bar{s}$ pointwise across every estimation window. We define the shrinkage counterpart $\hat{\boldsymbol{w}}_{\text{ef}}^{\text{LW}}$ analogously.

Table 1 summarizes the portfolio strategies evaluated in our numerical studies. Panel A reports the oracle combined portfolios, which are based on true optimal combination coefficients computed from the true $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$, as given in Propositions 3 and 6. These portfolios are feasible only in the simulation study and serve as theoretical upper bounds on the performance of the combined portfolios. Panel B summarizes the implementable combined portfolios constructed using four sets of estimators: the sample estimators in equation (14), the sophisticated estimators in equation (15), the shrinkage sample estimators in equation (22), and the shrinkage sophisticated estimators in equation (23). Panel C lists the plug-in portfolios of equation (2) and the ESG-constrained EF portfolios. Furthermore, Panel D outlines the ESG-unaware baseline

portfolios: the two-fund portfolio $\hat{\boldsymbol{w}}_k$ of Kan et al. (2022), and its shrinkage counterpart $\hat{\boldsymbol{w}}_k^{\mathrm{LW}}$. They do not utilize ESG scores in their portfolio construction.

[Insert Table 1 approximately here]

Finally, it is worth making a brief remark regarding the shrinkage-based strategies in Table 1. Although replacing the sample covariance matrix with its shrinkage counterpart in the constituent portfolios improves practical performance (Kan et al. (2022), Huang et al. (2024)), it renders the exact analytical derivation of the optimal combination coefficients and the expected OOS utility intractable. To circumvent this, we follow the pragmatic approach of Kan et al. (2022). We retain the functional forms of the combination coefficients derived under the sample covariance matrix (Propositions 3 and 6), and apply the shrinkage covariance matrix to both the constituent portfolios and the combination coefficients in the numerical implementation (Sets 3 and 4 in Panel B of Table 1). The same convention also applies to the portfolios $\hat{\boldsymbol{w}}_k^{\mathrm{LW}}$ and $\hat{\boldsymbol{w}}_{\mathrm{ef}}^{\mathrm{LW}}$. Furthermore, the mathematical properties regarding the realized portfolio factor exposure presented in Proposition 5 remain valid under this substitution.

B. Simulation Study

For the simulation study, we calibrate model parameters using empirical data from the constituents of the Dow Jones Industrial Average (DJ30) as of November 8, 2024. Specifically, we obtain annual ESG scores from Refinitiv[2] via Wharton Research Data Services (WRDS) spanning the ten-year period from 2014 to 2023. Because data for Honeywell International Inc. is unavailable via WRDS, we exclude this asset, thereby reducing the investment universe to

[2]We use Version 2 (V2) of the Refinitiv ESG Scores accessed via Wharton Research Data Services.

$N = 29$ stocks (listed in Appendix B.1). We compute the time-series average of the ESG scores for each asset, with summary statistics reported in Table 2. In our numerical experiments, we treat these historical averages as the deterministic factor vector $\boldsymbol{s}$ for portfolio construction.

[Insert Table 2 approximately here]

In addition to the ESG data, we obtain monthly adjusted closing prices for these 29 stocks from January 2014 to December 2023 to compute monthly returns. We calculate monthly excess returns by subtracting the one-month Treasury bill rate, sourced from Kenneth R. French's Data Library. From these excess returns, we compute the sample mean vector and sample covariance matrix, which we then treat as the true population parameters ($\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$) for the data-generating process in our simulation study.

To evaluate the OOS performance of implementable combined portfolios, we simulate $M = 10{,}000$ sets of $T \times N$ asset excess returns, each generated from the multivariate normal distribution parameterized by the calibrated population mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$. We set the estimation window length to $T = 120$. For each simulated set of $T \times N$ returns, we compute the sample mean and sample (or shrinkage) covariance matrix, and then use them to construct the corresponding implementable portfolios. The risk aversion coefficient is set to $\gamma = 5$.

Furthermore, we illustrate the selection of the target ESG exposure within our simulation framework. At the population parameters ($\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$), the unconstrained maximum expected OOS MV utility is $U_{\text{unc}} = 0.0318$, and the peak exposure is $\bar{s}^* = 0.2889$.[3] Table 3 reports the admissible intervals for $\bar{s}$ derived from equation (20) across various proportional thresholds α for

[3]Because short selling is permitted, the portfolio-level ESG score is not bounded by the minimum and maximum individual asset scores reported in Table 2.

the relative criterion in Proposition 7, alongside the corresponding maximum permissible utility concession $d = (1-\alpha)U_{\text{unc}}$ for the absolute criterion. While an ESG-motivated investor naturally selects the upper endpoint to maximize factor exposure, our analysis also reports the lower endpoints. Specifically, we evaluate both bounds for $\alpha \in \{0.98, 0.99\}$, and additional upper bounds for $\alpha \in \{0.80, 0.90, 0.95\}$. As established, the interval collapses precisely to $\bar{s}^*$ when $\alpha = 1$. Finally, note that the target levels in Table 3 are derived explicitly for the equality constraint. Under the inequality formulation, any target below $\bar{s}^*$ (the lower endpoints) renders the constraint inactive, allowing the portfolio to trivially attain the unconstrained maximum utility U_{unc}.

[Insert Table 3 approximately here]

1. The Portfolio ESG Scores

Because the constraints in the optimization problems in equations (8) and (16) target the expected portfolio ESG exposure, we evaluate the empirical distribution of the realized ESG scores in our Monte Carlo simulation. For any implementable portfolio $\tilde{\boldsymbol{w}}$, let $\tilde{\boldsymbol{w}}_m$ denote its realization in the m-th simulated sample for $m = 1, 2, \ldots, M$. We compute the cross-sample average and standard deviation of the portfolio ESG scores as follows:

$$\widehat{\text{ES}}_{\tilde{\boldsymbol{w}}} = \frac{1}{M}\sum_{m=1}^{M} \tilde{\boldsymbol{w}}_m^\top \boldsymbol{s} \qquad \text{and} \qquad \widehat{\text{VS}}_{\tilde{\boldsymbol{w}}} = \sqrt{\frac{1}{M-1}\sum_{m=1}^{M} (\tilde{\boldsymbol{w}}_m^\top \boldsymbol{s} - \widehat{\text{ES}}_{\tilde{\boldsymbol{w}}})^2}. \tag{25}$$

These quantities estimate the expectation $\mathbb{E}[\tilde{\boldsymbol{w}}^\top \boldsymbol{s}]$ and the standard deviation $\sqrt{\text{Var}[\tilde{\boldsymbol{w}}^\top \boldsymbol{s}]}$, respectively, and thus provide information on both the average level and the variability of portfolio ESG scores across simulated samples.

[Insert Table 4 approximately here]

Table 4 reports the average portfolio ESG scores and their standard deviations (in parentheses) across the target ESG levels $\bar{s}$ listed in Table 3. Panel A presents the expected values and standard deviations for the oracle combined portfolios, computed directly from the analytical expressions in Propositions 1 and 2 using the population parameters. Panel B reports the corresponding Monte Carlo estimates for the implementable portfolios using equation (25). The differences between Panels A and B primarily reflect estimation risk, as the implementable portfolios rely on estimated rather than true optimal combination coefficients. The portfolios from Panels C and D of Table 1 are omitted from this tabulation for two reasons. First, the plug-in portfolios (Panel C) satisfy the factor constraint pointwise by construction, and hence, their realized ESG scores equal the exact target level in every simulated sample. Second, the ESG-unaware portfolios (Panel D) are invariant to the target level. Specifically, the two-fund portfolios $\hat{\boldsymbol{w}}_k$ and $\hat{\boldsymbol{w}}_k^{\mathrm{LW}}$ achieve average ESG scores of 0.3671 (standard deviation 0.212) and 0.5308 (standard deviation 0.120), respectively.

From Table 4, we can draw three key observations regarding the variability and expected level of the portfolio exposures. First, under the equality-constrained formulation, the standard deviations reported in Panel B are substantially smaller than those in Panel A. This occurs because Panel A reflects the inherent variability of the oracle portfolio at the population level, whereas Panel B evaluates implementable portfolios constructed sample by sample. Specifically, the estimated combination coefficients force the realized portfolio ESG score to remain tightly anchored to the target, which reduces cross-sample dispersion of ESG scores. Under the inequality-constrained formulation, the standard deviations are also far below those in Panel A at

high target levels, where the constraint always binds. However, at low targets, the sample-dependent nature of the binding condition introduces an additional source of variation. For example, the standard deviation of $\tilde{\boldsymbol{w}}_{c\text{-}I}$ can exceed its oracle counterpart (0.199 versus 0.179 at $\bar{s} = 0.1090$).

Second, within each set in Panel B, the implementable equality-constrained portfolios (e.g., $\hat{\boldsymbol{w}}_{c\text{-}E}$ and $\tilde{\boldsymbol{w}}_{c\text{-}E}$) consistently exhibit a smaller standard deviation in their ESG scores than their inequality-constrained counterparts (e.g., $\hat{\boldsymbol{w}}_{c\text{-}I}$ and $\tilde{\boldsymbol{w}}_{c\text{-}I}$). Furthermore, as established in Proposition 5, the standard deviations for $\hat{\boldsymbol{w}}_{c\text{-}E}$ (Set 1) and $\hat{\boldsymbol{w}}_{c\text{-}E}^{\mathrm{LW}}$ (Set 3) are invariant to the target $\bar{s}$. Conversely, the portfolios based on the sophisticated estimators, $\tilde{\boldsymbol{w}}_{c\text{-}E}$ (Set 2) and $\tilde{\boldsymbol{w}}_{c\text{-}E}^{\mathrm{LW}}$ (Set 4), match the target exactly in every simulated sample, so their standard deviations are zero.

Third, Panel A of Table 4 reveals a clear shift in the ESG scores for the inequality-constrained portfolio $\hat{\boldsymbol{w}}_{c\text{-}I}^{*}$. The shift is governed by the sign of the parameter ζ, as discussed in Proposition 6, or equivalently, the relative magnitude of $\bar{s}$ to $\bar{s}^{*}$, as shown in Corollary 2. When the target $\bar{s}$ lies below the peak exposure $\bar{s}^{*} = 0.2889$, the inequality constraint is inactive, and the expected portfolio ESG score of $\hat{\boldsymbol{w}}_{c\text{-}I}^{*}$ remains constant. When the target reaches or exceeds the threshold ($\bar{s} \geq \bar{s}^{*}$), the constraint binds at the boundary and $\hat{\boldsymbol{w}}_{c\text{-}I}^{*}$ coincides with $\hat{\boldsymbol{w}}_{c\text{-}E}^{*}$, so their expected portfolio ESG scores match. In the oracle setting, ζ is known for any given target. By contrast, for the implementable combined portfolios, ζ must be estimated, and the estimation error can be severe enough to reverse its sign. Thus, whether the equality- and inequality-constrained implementable portfolios differ becomes a sample-dependent event.

In our simulation, the equality- and inequality-constrained implementable portfolios differ predominantly at low target levels and increasingly converge as the target rises. Table B.3 in Appendix B.2.1 corroborates this by reporting the proportion of the $M = 10{,}000$ simulated

samples where the estimated ζ is negative (i.e., where the two portfolios diverge). As the target $\bar{s}$ increases, this proportion decays toward zero. This convergence is also reflected in Panel B of Table 4, where the gap between the average ESG scores of the two portfolios narrows at higher targets. Furthermore, when the estimated ζ is negative, the implementable inequality-constrained portfolio $\hat{\boldsymbol{w}}_{c\text{-}I}$ generally delivers a higher ESG score than its equality-constrained counterpart $\hat{\boldsymbol{w}}_{c\text{-}E}$ (proof provided in Appendix A.12). Consequently, across all scenarios in Table 4, $\hat{\boldsymbol{w}}_{c\text{-}I}$ consistently achieves an average ESG score at least as high as that of $\hat{\boldsymbol{w}}_{c\text{-}E}$.

2. Out-of-Sample MV Utilities

We evaluate the various portfolios by examining their expected OOS MV utilities. For any implementable portfolio $\tilde{\boldsymbol{w}}$, the utility is estimated by averaging over the M simulated samples:

$$\widehat{\mathrm{EU}}_{\tilde{\boldsymbol{w}}} = \frac{1}{M}\sum_{m=1}^{M}\left(\tilde{\boldsymbol{w}}_m^\top \boldsymbol{\mu} - \frac{\gamma}{2}\,\tilde{\boldsymbol{w}}_m^\top \boldsymbol{\Sigma}\tilde{\boldsymbol{w}}_m\right), \tag{26}$$

where the cross-sample averages of $\tilde{\boldsymbol{w}}_m^\top \boldsymbol{\mu}$ and $\tilde{\boldsymbol{w}}_m^\top \boldsymbol{\Sigma}\tilde{\boldsymbol{w}}_m$ estimate the theoretical expected OOS mean $\mathbb{E}[\tilde{\boldsymbol{w}}^\top \boldsymbol{\mu}]$ and variance $\mathbb{E}[\tilde{\boldsymbol{w}}^\top \boldsymbol{\Sigma}\tilde{\boldsymbol{w}}]$, respectively.

Table 5 reports these estimated OOS MV utilities across the evaluated portfolios. For the oracle combined portfolios in Panel A, the expected OOS means and variances are computed directly from the analytical expressions in Proposition 1, which are then used to calculate the exact OOS MV utilities. For the implementable portfolios in Panels B and C, the expected utilities are estimated using the cross-sample average in equation (26). For completeness, the underlying expected OOS means and standard deviations are detailed in Appendix B.2.3.

[Insert Table 5 approximately here]

The ESG-unaware portfolios are omitted from Table 5 because their performance is invariant to the target ESG score. Under the population parameters, the two-fund portfolio of Kan et al. (2022) achieves an expected OOS MV utility of 0.0315, and yields a dominance interval of $(0.1578, 0.4200)$ for the target ESG score, as theoretically discussed in Corollary 1. This implies that, for any target $\bar{s}$ within this interval, the oracle equality-constrained portfolio $\hat{\boldsymbol{w}}^*_{c\text{-}E}$ strictly outperforms the two-fund benchmark. This interval maps to a maximum permissible utility concession of $d = 0.0003$, or equivalently, a proportional threshold of $\alpha \approx 0.99$, for the criteria established in Proposition 7.

This theoretical dominance largely extends to the implementable portfolios. The implementable two-fund portfolio yields an average utility of approximately 0.03 under both the sample and shrinkage covariance matrices. For any $\bar{s}$ within the dominance interval, our combined three-fund portfolios (Sets 1–3) deliver comparable or strictly superior performance. The relative underperformance of Set 4 likely occurs because the sophisticated estimators are explicitly designed to correct sample bias. Applying them to the shrinkage covariance matrix, which is already inherently bias-corrected, results in excessive over-shrinkage of the combination coefficients.

We draw three key observations from Table 5. First, in Panel A, the utilities of the two oracle portfolios vary with the target $\bar{s}$ in exact accordance with our theoretical predictions (see also Figure B.1 in Appendix B.2.2). At the peak exposure $\bar{s}^* = 0.2889$, both portfolios attain their maximum expected OOS utility. When $\bar{s} < \bar{s}^*$, the inequality constraint is inactive, so the utility of $\hat{\boldsymbol{w}}^*_{c\text{-}I}$ remains flat at its unconstrained maximum, whereas $\hat{\boldsymbol{w}}^*_{c\text{-}E}$ is forced to adhere to the equality constraint and delivers a lower utility. When $\bar{s} \geq \bar{s}^*$, the inequality constraint binds, and

thus the expected utilities of $\hat{\boldsymbol{w}}^*_{c\text{-}I}$ and $\hat{\boldsymbol{w}}^*_{c\text{-}E}$ are the same. This confirms that forcing the target ESG score beyond the unconstrained peak entails a utility reduction.

Second, the implementable combined portfolios in Panel B generally exhibit lower utilities than their oracle counterparts in Panel A, yet they substantially outperform the benchmark portfolios in Panel C. Rare exceptions arise at the target levels farthest from $\bar{s}^*$, where the implementable portfolios in Sets 1 and 3 actually attain slightly higher utilities than the oracle. This counterintuitive result is likely driven by two factors. Primarily, at extreme targets (e.g., $\bar{s} = 0.8579$), the realized ESG scores of these implementable portfolios fall short of the target and lie closer to the optimum $\bar{s}^*$ (e.g., 0.8139 for Set 1 and 0.8188 for Set 3, as seen in Table 4). Because the utility cost is quadratic in the distance from $\bar{s}^*$, this estimation error paradoxically improves utility. Furthermore, Set 3 applies the shrinkage covariance matrix to its constituent portfolios, which organically reduces estimation error and pushes its utilities above Set 1 across all target levels, making these exceptions more frequent. Broadly, the proposed three-fund framework successfully prevents the catastrophic OOS deterioration seen in the benchmark approaches, such as $\hat{\boldsymbol{w}}_{\text{ef}}$, whose utilities collapse to the order of -10^3.

Third, within Panel B, the implementable equality- and inequality-constrained portfolios ($\hat{\boldsymbol{w}}_{c\text{-}E}$ and $\hat{\boldsymbol{w}}_{c\text{-}I}$) exhibit distinct utility dynamics. As discussed in Section V.B.1, these differences manifest in simulated samples where the estimated parameter ζ is negative. We detail the exact mathematical difference in OOS utility between these two formulations in Appendix A.12. Because this difference depends on a complex interaction of parameters (including the risk aversion coefficient, the number of assets, and the estimation window length), neither portfolio uniformly dominates the other across all scenarios. To quantify this numerically, Table B.4 in Appendix B.2.1 reports the proportion of simulated samples where $\hat{\boldsymbol{w}}_{c\text{-}E}$ attains a higher OOS

MV utility than $\hat{\boldsymbol{w}}_{c\text{-}I}$, conditional on the two portfolios differing from each other. The results demonstrate that $\hat{\boldsymbol{w}}_{c\text{-}E}$ delivers a higher utility in the majority of these scenarios, particularly at relatively high target ESG scores. Consequently, the average OOS utility of the implementable $\hat{\boldsymbol{w}}_{c\text{-}E}$ generally exceeds that of $\hat{\boldsymbol{w}}_{c\text{-}I}$, as observed directly in Panel B of Table 5.

Additionally, we examine alternative risk aversion coefficients ($\gamma \in \{3, 10\}$) and estimation windows ($T \in \{60, 240\}$), and report the corresponding utilities in Appendices B.2.5 and B.2.6, respectively. While the equality-constrained formulation dominates for $\gamma = 3$ and 5, the inequality constraint regains its theoretical advantage at nonbinding target levels when the risk aversion coefficient is extremely high ($\gamma = 10$). The estimation window, in contrast, affects the absolute level of the utilities but preserves the relative ranking of the estimator sets in the majority of cases. A shorter window ($T = 60$) amplifies estimation risk, thereby lowering the utilities of all implementable combined portfolios (with Set 3 delivering the most robust performance). Conversely, a longer window ($T = 240$) mitigates this risk, raising their utilities and causing the performance of the four estimator sets to largely converge.

C. Empirical Study

While our simulation study relies on 29 constituents of the Dow Jones Industrial Average, our empirical analysis evaluates portfolio performance across a much broader investment universe. Specifically, we consider up to 100 major stocks selected from the S&P 500 (listed in Appendix B.1). To construct this universe, we rank the S&P 500 constituents by their market capitalization as of 2024 and select the top 100 stocks that possess complete historical price and ESG data. We obtain annually updated ESG scores from Refinitiv via WRDS covering the period

from 2013 to 2023. To prevent look-ahead bias, we apply the ESG score updated at the end of calendar year $t-1$ to all portfolio construction throughout the subsequent investment year t. Accordingly, we focus the investment analysis on the period from 2014-01-01 to 2024-12-31.

We obtain weekly adjusted closing prices for these stocks and calculate weekly returns as the relative price changes between consecutive weeks. The corresponding excess returns are calculated by subtracting the weekly Treasury bill rate, sourced from Kenneth R. French's data library. The use of weekly data strikes a balance between data availability and estimation reliability. Compared to monthly data, weekly observations provide a larger effective sample size for estimating portfolio rules and evaluating OOS performance, while remaining less susceptible to market microstructure noise and short-term trading frictions than daily data. Although requiring a complete data history facilitates implementation, it inherently introduces survivorship bias. However, because our primary objective is the relative comparison of different portfolios evaluated on an identical dataset, we expect that this bias will not materially affect our main conclusions.

We employ a rolling-window approach with an estimation window of $T = 120$ weeks. Starting from week $T+1$, we use the preceding T weeks of returns to compute the necessary sample mean vectors and sample (or shrinkage) covariance matrices, which are then used to construct the implementable portfolios for the subsequent week. To form the estimation window for the first investment week in 2014, we incorporate historical returns dating back to 2011-09-19. More generally, if X denotes the total number of weeks in the sample period, this rolling-window procedure yields $X - T$ realized OOS weekly returns for each portfolio. Results utilizing alternative window lengths ($T \in \{60, 240\}$) are reported in Appendix B.3.4.

We construct implementable portfolios across various universe sizes: $N \in \{6, 10, 15, 25,$

$35, 55, 75, 100\}$, where each case uses the top N stocks listed in Table B.2. The risk aversion coefficient is fixed at $\gamma = 5$. We present results utilizing the linear shrinkage covariance matrix, as it delivers the highest certainty equivalent returns in nearly every scenario. Corresponding results using the sample and nonlinear shrinkage covariance matrices are deferred to Appendix B.3.2. Furthermore, results for $\gamma \in \{3, 10\}$ are provided in Appendix B.3.3.

We begin by examining the empirical distribution of the ESG scores for the implementable combined portfolios. For each strategy, we compute the cross-window average and standard deviation of its realized ESG scores. Table 6 reports these results across three target ESG levels: $\bar{s} \in \{0.5630, 0.7018, 0.8059\}$, selected via the relative criterion established in Proposition 7. Because each asset universe size N defines its own unconstrained peak $\bar{s}^*$ (and consequently, its own implied target levels), we anchor our target selection at $N = 25$ to ensure a standardized comparison across all universe sizes. Using the full-sample moments of the asset returns (2011-09-19 to 2024-12-31) and the time-series average of the annual ESG scores (2013–2023) as fixed inputs, we obtain these three targets as the upper endpoints of the interval defined in equation (20) evaluated at $\alpha \in \{0.99, 0.95, 0.90\}$.

[Insert Table 6 approximately here]

We draw three key observations from Table 6. First, the average ESG scores of the implementable equality-constrained portfolios ($\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$ and $\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$) are consistently lower than those of their inequality-constrained counterparts ($\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$ and $\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$), except for a single tie at $N = 100$ when $\bar{s} = 0.8059$. This aligns with theoretical expectations, as the inequality formulation allows the portfolio to exceed the target $\bar{s}$. Second, the impact of estimation error on the combination coefficients becomes increasingly pronounced as the asset universe size N expands. For relatively

small universes (e.g., $N \in \{6, 10\}$), the average ESG scores of $\hat{\boldsymbol{w}}_{c\text{-}E}^{\mathrm{LW}}$ remain tightly anchored to the target $\bar{s}$. Conversely, for larger asset universe sizes (e.g., $N \in \{75, 100\}$), heightened estimation risk causes a wider divergence between the realized average score and the target. Third, across the vast majority of scenarios, the average ESG scores of $\hat{\boldsymbol{w}}_{c\text{-}I}^{\mathrm{LW}}$ and $\tilde{\boldsymbol{w}}_{c\text{-}I}^{\mathrm{LW}}$ strictly exceed the target $\bar{s}$.

To evaluate empirical performance, we calculate the sample mean ($\hat{\mu}_{\tilde{\boldsymbol{w}}}$) and sample variance ($\hat{\sigma}^2_{\tilde{\boldsymbol{w}}}$) of the OOS weekly excess returns over the $X - T$ evaluation weeks for each implementable portfolio $\tilde{\boldsymbol{w}}$. We annualize these quantities by multiplying them by 52. The annualized certainty equivalent return (CER) is then computed as:

$$\mathrm{CER}_{\tilde{\boldsymbol{w}}} = \hat{\mu}_{\tilde{\boldsymbol{w}}} - \frac{\gamma}{2}\hat{\sigma}^2_{\tilde{\boldsymbol{w}}}.$$

The CER represents the guaranteed return, in excess of the risk-free rate, that an investor would accept in lieu of taking on the risk associated with the portfolio $\tilde{\boldsymbol{w}}$. Table 7 reports these annualized CERs across all evaluated strategies. For completeness, the underlying annualized average OOS returns and standard deviations are detailed in Appendix B.3.1.

[Insert Table 7 approximately here]

Several key observations emerge from Table 7. First, the implementable combined portfolios consistently outperform the benchmark portfolios $\hat{\boldsymbol{w}}_{\mathrm{mv}}^{\mathrm{LW}}$ and $\hat{\boldsymbol{w}}_{\mathrm{ef}}^{\mathrm{LW}}$ across all asset universe sizes and target ESG scores. While the benchmark $\hat{\boldsymbol{w}}_{\mathrm{mv}}^{\mathrm{LW}}$ generally delivers higher average returns, it is plagued by substantially larger standard deviations (see Appendix B.3.1). In some cases, although the average return of $\hat{\boldsymbol{w}}_{\mathrm{mv}}^{\mathrm{LW}}$ is roughly double that of the combined

portfolios, its standard deviation is more than three times as large. This contrast becomes even more pronounced as the number of assets increases. This suggests that the higher average returns of $\hat{\boldsymbol{w}}_{\text{mv}}^{\text{LW}}$ are more than offset by the accompanying spike in volatility, which leads to a severe deterioration in CER.

Second, comparing the portfolios in Table 7 reveals that those based on the sophisticated estimators outperform their sample counterparts for asset universe sizes up to $N = 55$, with a marginal exception at $\bar{s} = 0.8059$, whereas the ordering generally reverses at $N = 75$ and 100. As discussed in Section V.B.2, applying sophisticated bias corrections to an already-shrunk covariance matrix results in excessive over-shrinkage. This empirical analysis reveals that this detrimental double-adjustment effect becomes increasingly severe as the asset universe size N grows relative to the estimation window T. By contrast, when the standard sample covariance matrix is used, the sophisticated estimators strictly deliver higher CERs than their sample-based counterparts at every asset universe size, as reported in Table B.11 of Appendix B.3.2.

Third, the performance comparison between the two constraint formulations separates cleanly by asset universe size. The inequality-constrained portfolio $\tilde{\boldsymbol{w}}_{c\text{-}I}^{\text{LW}}$ attains higher annualized CERs at every target level for $N \leq 15$, whereas the ordering reverses in favor of the equality-constrained version $\tilde{\boldsymbol{w}}_{c\text{-}E}^{\text{LW}}$ for $N \geq 25$. In theory, under the multivariate normality assumption and in the absence of estimation risk in the combination coefficients, the optimal combined portfolio $\hat{\boldsymbol{w}}_{c\text{-}I}^{*}$ should never underperform $\hat{\boldsymbol{w}}_{c\text{-}E}^{*}$. The empirical deviation from this theoretical result likely arises from two factors. First, real-world asset returns often exhibit heavy tails or asymmetries that violate the multivariate normality assumption. Second, as we previously discussed, the inequality-constrained portfolio relies on the estimated parameter ζ to determine the binding status of the constraint. The adverse impact of this specific estimation risk scales with

dimensionality, which explains why the flexibility of the inequality constraint pays off only at smaller asset universe sizes where the estimation error in the combination coefficients remains less severe.

[Insert Table 8 approximately here]

For comparison, we also evaluate the performance of the ESG-unaware two-fund rules of Kan et al. (2022), as reported in Table 8. Comparing these benchmarks against our proposed ESG-constrained three-fund framework (Tables 6 and 7) highlights the distinct advantages of our methodology in delivering both reliable sustainability control and superior OOS financial performance. First, while the ESG-unaware rules ($\hat{\boldsymbol{w}}_k$ and $\hat{\boldsymbol{w}}_k^{\mathrm{LW}}$) attempt to optimize unconstrained MV efficiency, their realized average ESG scores (Panel A of Table 8) are entirely uncontrolled and fluctuate substantially across varying universe sizes N. In contrast, our proposed method successfully anchors the portfolio at or near the investor's target exposure $\bar{s}$. Second, because the ESG-unaware rules passively realize their average ESG scores without any optimization control, we can evaluate our proposed methods at prespecified target levels ($\bar{s}$) that neatly bracket these passively realized scores. We focus this comparison on $\hat{\boldsymbol{w}}_k^{\mathrm{LW}}$ since it yields strictly better CERs than $\hat{\boldsymbol{w}}_k$. We evaluate it against our equality-constrained shrinkage portfolios ($\hat{\boldsymbol{w}}_{c\text{-}E}^{\mathrm{LW}}$ and $\tilde{\boldsymbol{w}}_{c\text{-}E}^{\mathrm{LW}}$). As detailed in Panel C of Table 8, when evaluated at these bracketing target levels, our portfolio $\tilde{\boldsymbol{w}}_{c\text{-}E}^{\mathrm{LW}}$ achieves higher annualized CERs than $\hat{\boldsymbol{w}}_k^{\mathrm{LW}}$ at every asset universe size, except at the upper bracketing target for $N = 75$ and 100, while the portfolio $\hat{\boldsymbol{w}}_{c\text{-}E}^{\mathrm{LW}}$ falls short of the benchmark in only a few cells.

VI. Conclusion

This paper develops a constrained portfolio combination approach that mitigates estimation risk while satisfying explicit factor exposure mandates. Building upon the three-fund separation for the population-optimal portfolio, we design an implementable combined three-fund rule that inherently preserves the full-investment constraint. Through the maximization of expected OOS MV utility, we derive analytical combination coefficients under both equality and inequality expected factor constraints. A central contribution of this framework is that the proposed sophisticated estimators allow the implementable portfolio to satisfy the target factor exposure pointwise in every realization, which effectively overcomes the operational infeasibility of strict constraints under parameter uncertainty. Furthermore, we provide explicit utility-based criteria to guide investors in selecting target factor levels.

Our simulation and empirical studies, illustrated through the lens of ESG investing, confirm the practical efficacy of this methodology. The proposed combined portfolios consistently deliver substantial improvements in expected OOS utility relative to standard plug-in and efficient-frontier benchmarks. Importantly, our analysis reveals a structural trade-off in constraint formulation. By avoiding the estimation noise associated with determining constraint boundaries, the equality-constrained portfolio typically achieves superior OOS utility, even at target levels where the constraint is theoretically nonbinding. The theoretical flexibility of the inequality constraint translates into an empirical advantage only in smaller asset universes, or when the risk aversion coefficient in the MV framework is extremely high. Ultimately, while demonstrated using ESG scores, this paper provides an analytical framework for incorporating any quantitative characteristic mandate into MV optimal portfolio management.

References

Avramov, D.; S. Cheng; A. Lioui; and A. Tarelli. "Sustainable Investing with ESG Rating Uncertainty." *Journal of Financial Economics*, 145 (2022), 642–664.

Berg, F.; J. F. Koelbel; and R. Rigobon. "Aggregate Confusion: The Divergence of ESG Ratings." *Review of Finance*, 26 (2022), 1315–1344.

Bodnar, T.; Y. Okhrin; and N. Parolya. "Optimal Shrinkage-Based Portfolio Selection in High Dimensions." *Journal of Business & Economic Statistics*, 41 (2023), 140–156.

Chavez-Bedoya, L., and R. Castaneda. "Formulas for Calculating Out-of-Sample Performance of Constrained Portfolios." Working Paper, Universidad ESAN (2024).

DeMiguel, V.; L. Garlappi; F. J. Nogales; and R. Uppal. "A Generalized Approach to Portfolio Optimization: Improving Performance by Constraining Portfolio Norms." *Management Science*, 55 (2009), 798–812.

DeMiguel, V.; A. Martín-Utrera; and F. J. Nogales. "Size Matters: Optimal Calibration of Shrinkage Estimators for Portfolio Selection." *Journal of Banking & Finance*, 37 (2013), 3018–3034.

DeMiguel, V.; A. Martín-Utrera; and F. J. Nogales. "Parameter Uncertainty in Multiperiod Portfolio Optimization with Transaction Costs." *Journal of Financial and Quantitative Analysis*, 50 (2015), 1443–1471.

Gupta, A. K., and D. K. Nagar. *Matrix Variate Distributions*. Boca Raton, FL: Chapman and Hall/CRC (2018).

Huang, Z.; P. Wei; and C. Weng. "Tail Mean-Variance Portfolio Selection with Estimation Risk." *Insurance: Mathematics and Economics*, 116 (2024), 218–234.

Jagannathan, R., and T. Ma. "Risk Reduction in Large Portfolios: Why Imposing the Wrong Constraints Helps." *Journal of Finance*, 58 (2003), 1651–1683.

Jorion, P. "Bayes-Stein Estimation for Portfolio Analysis." *Journal of Financial and Quantitative Analysis*, 21 (1986), 279–292.

Jorion, P. "Bayesian and CAPM Estimators of the Means: Implications for Portfolio Selection." *Journal of Banking & Finance*, 15 (1991), 717–727.

Kalymon, B. A. "Estimation Risk in the Portfolio Selection Model." *Journal of Financial and Quantitative Analysis*, 6 (1971), 559–582.

Kan, R., and N. Lassance. "Optimal Portfolio Choice with Fat Tails and Parameter Uncertainty." *Journal of Financial and Quantitative Analysis*, 60 (2025), 3753–3790.

Kan, R.; X. Wang; and G. Zhou. "Optimal Portfolio Choice with Estimation Risk: No Risk-Free Asset Case." *Management Science*, 68 (2022), 2047–2068.

Kan, R., and G. Zhou. "Optimal Portfolio Choice with Parameter Uncertainty." *Journal of Financial and Quantitative Analysis*, 42 (2007), 621–656.

Kircher, F., and D. Rösch. "A Shrinkage Approach for Sharpe Ratio Optimal Portfolios with Estimation Risks." *Journal of Banking & Finance*, 133 (2021), 106281.

Klein, R. W., and V. S. Bawa. "The Effect of Estimation Risk on Optimal Portfolio Choice." *Journal of Financial Economics*, 3 (1976), 215–231.

Klein, R. W., and V. S. Bawa. "The Effect of Limited Information and Estimation Risk on Optimal Portfolio Diversification." *Journal of Financial Economics*, 5 (1977), 89–111.

Lassance, N. "Reconciling Mean-Variance Portfolio Theory with Non-Gaussian Returns." *European Journal of Operational Research*, 297 (2022), 729–740.

Lassance, N.; A. Martín-Utrera; and M. Simaan. "The Risk of Expected Utility under Parameter Uncertainty." *Management Science*, 70 (2024a), 7644–7663.

Lassance, N.; R. Vanderveken; and F. Vrins. "On the Combination of Naive and Mean-Variance Portfolio Strategies." *Journal of Business & Economic Statistics*, 42 (2024b), 875–889.

Ledoit, O., and M. Wolf. "A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices." *Journal of Multivariate Analysis*, 88 (2004), 365–411.

Ledoit, O., and M. Wolf. "Nonlinear Shrinkage of the Covariance Matrix for Portfolio Selection: Markowitz Meets Goldilocks." *Review of Financial Studies*, 30 (2017), 4349–4388.

Ledoit, O., and M. Wolf. "Analytical Nonlinear Shrinkage of Large-Dimensional Covariance Matrices." *Annals of Statistics*, 48 (2020), 3043–3065.

Lo, A. W., and R. Zhang. "Performance Attribution for Portfolio Constraints." *Management Science*, 71 (2025), 7537–7559.

Luo, D.; X. Shan; J. Yan; and Q. Yan. "Sustainable Investment under ESG Volatility and Ambiguity." *Economic Modelling*, 128 (2023), 106471.

Markowitz, H. "Portfolio Selection." *Journal of Finance*, 7 (1952), 77–91.

Muirhead, R. J. *Aspects of Multivariate Statistical Theory*. New York, NY: John Wiley & Sons (1982).

Okhrin, Y., and W. Schmid. "Distributional Properties of Portfolio Weights." *Journal of Econometrics*, 134 (2006), 235–256.

Pedersen, L. H.; S. Fitzgibbons; and L. Pomorski. "Responsible Investing: The ESG-Efficient Frontier." *Journal of Financial Economics*, 142 (2021), 572–597.

Simaan, Y. "The Opportunity Cost of Mean–Variance Choice under Estimation Risk." *European Journal of Operational Research*, 234 (2014), 382–391.

Tu, J., and G. Zhou. "Markowitz Meets Talmud: A Combination of Sophisticated and Naive Diversification Strategies." *Journal of Financial Economics*, 99 (2011), 204–215.

Wu, R.; Y. Lu; and Y. Yang. "Sustainable Investment: ESG Impacts on Large Portfolio." Working Paper, available at arXiv:2602.14439 (2026).

TABLE 1

List of Portfolio Strategies

Panel A lists the oracle combined portfolios, and Panel B lists their implementable versions constructed from different sets of estimators. Panels C and D list the benchmark portfolios and the ESG-unaware portfolios, respectively. ESG stands for environmental, social, and governance. The subscripts $c\text{-}E$, $c\text{-}I$, mv, ef, and k indicate the equality-constrained combined portfolio, the inequality-constrained combined portfolio, the mean-variance portfolio, the efficient-frontier portfolio, and the two-fund portfolio of Kan et al. (2022), respectively. The superscript LW indicates the use of the linear shrinkage covariance matrix of Ledoit and Wolf (2004).

Notation	Description
Panel A: Oracle Combined Portfolios (Simulation Only)	
$\hat{\boldsymbol{w}}^*_{c\text{-}E}$ and $\hat{\boldsymbol{w}}^*_{c\text{-}I}$	optimal combined portfolios in equations (10) and (18), respectively
Panel B: Implementable Versions of the Combined Portfolios $\hat{\boldsymbol{w}}^*_{c\text{-}E}$ and $\hat{\boldsymbol{w}}^*_{c\text{-}I}$	
$\hat{\boldsymbol{w}}_{c\text{-}E}$ and $\hat{\boldsymbol{w}}_{c\text{-}I}$	Set 1: apply sample estimators in equation (14)
$\tilde{\boldsymbol{w}}_{c\text{-}E}$ and $\tilde{\boldsymbol{w}}_{c\text{-}I}$	Set 2: apply sophisticated estimators in equation (15)
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$ and $\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	Set 3: apply shrinkage sample estimators in equation (22)
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$ and $\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	Set 4: apply shrinkage sophisticated estimators in equation (23)
Panel C: Benchmark Portfolios	
$\hat{\boldsymbol{w}}_{\mathrm{mv}}$ and $\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	plug-in portfolios for equation (2) based on sample and shrinkage covariance matrices
$\hat{\boldsymbol{w}}_{\mathrm{ef}}$ and $\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	plug-in portfolios for equation (24) based on sample and shrinkage covariance matrices
Panel D: ESG-Unaware Portfolios	
$\hat{\boldsymbol{w}}_k$ and $\hat{\boldsymbol{w}}^{\mathrm{LW}}_k$	plug-in portfolios for $\hat{\boldsymbol{w}}^*_k$ of Kan et al. (2022) based on sample and shrinkage covariance matrices, respectively

TABLE 2

Descriptive Statistics of the ESG Scores of Components in DJ30

This table reports descriptive statistics of the environmental, social, and governance (ESG) scores of the 29 constituents of the Dow Jones Industrial Average (DJ30) used in the simulation study. The ESG score of each stock is the time-series average of its annual Refinitiv ESG scores from 2014 to 2023.

Average	Smallest	25% Quantile	Median	75% Quantile	Largest
0.7336	0.5623	0.6868	0.7330	0.7682	0.9178

TABLE 3

Target ESG Levels and Utility Concessions Implied by the Fraction α of U_{unc}

For each value of α, this table reports the corresponding target environmental, social, and governance (ESG) level $\bar{s}$, given by the endpoints of the interval in equation (20), and the corresponding utility concession $d = (1-\alpha)U_{\mathrm{unc}}$. The quantity U_{unc} denotes the maximum expected out-of-sample mean-variance utility attainable by the combined portfolio without the factor constraint, evaluated at the population parameters of the simulation study, and $\bar{s}^*$ denotes the peak exposure, that is, the target ESG level at which the equality-constrained combined portfolio attains this maximum utility. The first two columns are the lower endpoints for $\alpha = 0.98, 0.99$, the third column is $\bar{s}^*$ attained at $\alpha = 1$, and the remaining columns are the upper endpoints for $\alpha = 0.99, 0.98, 0.95, 0.9, 0.8$.

	Lower Endpoint		$\bar{s}^*$	Upper Endpoint				
α	0.98	0.99	1.00	0.99	0.98	0.95	0.90	0.80
$\bar{s}$	0.1090	0.1617	0.2889	0.4161	0.4688	0.5734	0.6913	0.8579
d	0.0006	0.0003	0	0.0003	0.0006	0.0016	0.0032	0.0064

TABLE 4

The Average Portfolio ESG Scores in the Simulation Study

This table reports the average portfolio environmental, social, and governance (ESG) scores and their standard deviations (in parentheses) for different target ESG levels $\bar{s}$. The subscripts c-E and c-I indicate the equality- and inequality-constrained combined portfolios, respectively, and the superscript LW indicates the use of the linear shrinkage covariance matrix of Ledoit and Wolf (2004). The differences between Panels A and B reflect the estimation error in the two combination coefficients. Consistent with Proposition 5, the standard deviations of the ESG scores of $\hat{\boldsymbol{w}}_{c\text{-}E}$ and $\hat{\boldsymbol{w}}_{c\text{-}E}^{\text{LW}}$ in Sets 1 and 3 are invariant to the target level, while $\tilde{\boldsymbol{w}}_{c\text{-}E}$ and $\tilde{\boldsymbol{w}}_{c\text{-}E}^{\text{LW}}$ in Sets 2 and 4 attain the ESG target pointwise, with zero standard deviation.

$\bar{s}$	0.1090	0.1617	0.2889	0.4161	0.4688	0.5734	0.6913	0.8579
Panel A: Oracle Combined Portfolio								
$\hat{\boldsymbol{w}}_{c\text{-}E}^{*}$	0.1090	0.1617	0.2889	0.4161	0.4688	0.5734	0.6913	0.8579
	(0.189)	(0.186)	(0.179)	(0.173)	(0.172)	(0.169)	(0.169)	(0.171)
$\hat{\boldsymbol{w}}_{c\text{-}I}^{*}$	0.2889	0.2889	0.2889	0.4161	0.4688	0.5734	0.6913	0.8579
	(0.179)	(0.179)	(0.179)	(0.173)	(0.172)	(0.169)	(0.169)	(0.171)
Panel B: Implementable Combined Portfolio								
Set 1: Apply the Sample Estimators in Equation (14)								
$\hat{\boldsymbol{w}}_{c\text{-}E}$	0.2522	0.2917	0.3871	0.4825	0.5220	0.6005	0.6889	0.8139
	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)	(0.011)
$\hat{\boldsymbol{w}}_{c\text{-}I}$	0.3718	0.3901	0.4446	0.5129	0.5447	0.6125	0.6944	0.8154
	(0.158)	(0.144)	(0.112)	(0.081)	(0.070)	(0.051)	(0.034)	(0.020)
Set 2: Apply the Sophisticated Estimators in Equation (15)								
$\tilde{\boldsymbol{w}}_{c\text{-}E}$	0.1090	0.1617	0.2889	0.4161	0.4688	0.5734	0.6913	0.8579
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
$\tilde{\boldsymbol{w}}_{c\text{-}I}$	0.3266	0.3401	0.3894	0.4645	0.5028	0.5888	0.6969	0.8589
	(0.199)	(0.185)	(0.145)	(0.102)	(0.085)	(0.056)	(0.032)	(0.013)
Set 3: Apply the Shrinkage Sample Estimators in Equation (22)								
$\hat{\boldsymbol{w}}_{c\text{-}E}^{\text{LW}}$	0.2571	0.2966	0.3920	0.4874	0.5270	0.6054	0.6938	0.8188
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
$\hat{\boldsymbol{w}}_{c\text{-}I}^{\text{LW}}$	0.4813	0.4850	0.5032	0.5417	0.5646	0.6215	0.6990	0.8194
	(0.142)	(0.136)	(0.115)	(0.085)	(0.072)	(0.047)	(0.026)	(0.011)
Set 4: Apply the Shrinkage Sophisticated Estimators in Equation (23)								
$\tilde{\boldsymbol{w}}_{c\text{-}E}^{\text{LW}}$	0.1090	0.1617	0.2889	0.4161	0.4688	0.5734	0.6913	0.8579
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
$\tilde{\boldsymbol{w}}_{c\text{-}I}^{\text{LW}}$	0.4762	0.4771	0.4844	0.5112	0.5328	0.5970	0.6968	0.8582
	(0.151)	(0.149)	(0.137)	(0.108)	(0.091)	(0.056)	(0.026)	(0.005)

TABLE 5

Expected OOS MV Utilities for Portfolios in the Simulation Study

This table reports the expected out-of-sample (OOS) mean-variance (MV) utilities for different target factor levels $\bar{s}$. Panels A–C present the results for the portfolios listed in Table 1. The subscripts c-E, c-I, mv, and ef indicate the equality-constrained combined, inequality-constrained combined, plug-in MV, and efficient-frontier portfolios, respectively. The superscript LW indicates the use of the linear shrinkage covariance matrix of Ledoit and Wolf (2004), whereas portfolios without it use the sample covariance matrix. Entries below "-10^4" are denoted by "—". The proposed combined portfolios consistently outperform the benchmark portfolios in terms of utilities.

$\bar{s}$	0.1090	0.1617	0.2889	0.4161	0.4688	0.5734	0.6913	0.8579
Panel A: Oracle Combined Portfolio								
$\hat{\boldsymbol{w}}^{*}_{c\text{-}E}$	0.0312	0.0315	0.0318	0.0315	0.0312	0.0302	0.0286	0.0254
$\hat{\boldsymbol{w}}^{*}_{c\text{-}I}$	0.0318	0.0318	0.0318	0.0315	0.0312	0.0302	0.0286	0.0254
Panel B: Implementable Combined Portfolio								
Set 1: Apply the Sample Estimators in Equation (14)								
$\hat{\boldsymbol{w}}_{c\text{-}E}$	0.0312	0.0312	0.0310	0.0304	0.0301	0.0292	0.0280	0.0257
$\hat{\boldsymbol{w}}_{c\text{-}I}$	0.0305	0.0305	0.0304	0.0300	0.0297	0.0290	0.0279	0.0257
Set 2: Apply the Sophisticated Estimators in Equation (15)								
$\tilde{\boldsymbol{w}}_{c\text{-}E}$	0.0306	0.0309	0.0311	0.0308	0.0304	0.0295	0.0279	0.0247
$\tilde{\boldsymbol{w}}_{c\text{-}I}$	0.0303	0.0304	0.0305	0.0303	0.0300	0.0292	0.0278	0.0247
Set 3: Apply the Shrinkage Sample Estimators in Equation (22)								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.0318	0.0318	0.0316	0.0311	0.0308	0.0301	0.0289	0.0268
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.0308	0.0308	0.0308	0.0306	0.0304	0.0298	0.0288	0.0268
Set 4: Apply the Shrinkage Sophisticated Estimators in Equation (23)								
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.0277	0.0280	0.0283	0.0281	0.0278	0.0270	0.0255	0.0227
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.0274	0.0274	0.0274	0.0274	0.0272	0.0267	0.0255	0.0227
Panel C: Benchmark Portfolio								
$\hat{\boldsymbol{w}}_{\mathrm{mv}}$	-0.0268	-0.0265	-0.0263	-0.0266	-0.0269	-0.0279	-0.0295	-0.0326
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	0.0122	0.0125	0.0126	0.0122	0.0118	0.0108	0.0092	0.0061
$\hat{\boldsymbol{w}}_{\mathrm{ef}}$	—	—	-7806	-3782	-2538	-800.7	-12.2	-1011
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	-2072	-1750	-1086	-579.1	-415.5	-171.1	-23.73	-46.94

TABLE 6

The Average Portfolio ESG Scores for Portfolios in the Empirical Study

This table reports the average portfolio environmental, social, and governance (ESG) scores across rolling windows, with standard deviations in parentheses. N denotes the number of stocks; and $\bar{s}$ denotes the target ESG level. The subscripts c-E and c-I indicate the equality- and inequality-constrained combined portfolios, respectively; a hat and a tilde indicate that the combination coefficients are computed with the sample and sophisticated estimators, respectively; and the superscript LW indicates the use of the linear shrinkage covariance matrix of Ledoit and Wolf (2004).

N	6	10	15	25	35	55	75	100
Panel A: $\bar{s} = 0.5630$								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.5742	0.5665	0.5678	0.5868	0.5897	0.6131	0.6224	0.6288
	(0.005)	(0.007)	(0.014)	(0.020)	(0.023)	(0.049)	(0.051)	(0.058)
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.5630	0.5630	0.5630	0.5630	0.5630	0.5630	0.5630	0.5630
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.7844	0.8390	0.8266	0.7723	0.6979	0.6890	0.6483	0.6350
	(0.187)	(0.335)	(0.361)	(0.236)	(0.182)	(0.138)	(0.082)	(0.063)
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.7829	0.8391	0.8291	0.7648	0.6877	0.6803	0.6299	0.6256
	(0.189)	(0.335)	(0.360)	(0.242)	(0.188)	(0.141)	(0.085)	(0.063)
Panel B: $\bar{s} = 0.7018$								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.7049	0.6926	0.6880	0.6955	0.6869	0.6872	0.6733	0.6508
	(0.005)	(0.007)	(0.014)	(0.020)	(0.023)	(0.049)	(0.051)	(0.058)
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.7018	0.7018	0.7018	0.7018	0.7018	0.7018	0.7018	0.7018
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.8265	0.8991	0.8833	0.8251	0.7588	0.7319	0.6856	0.6520
	(0.144)	(0.289)	(0.323)	(0.197)	(0.147)	(0.109)	(0.067)	(0.059)
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.8273	0.9060	0.8942	0.8313	0.7743	0.7540	0.7169	0.7063
	(0.143)	(0.283)	(0.316)	(0.192)	(0.137)	(0.089)	(0.036)	(0.011)
Panel C: $\bar{s} = 0.8059$								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.8029	0.7872	0.7783	0.7771	0.7597	0.7427	0.7114	0.6673
	(0.005)	(0.007)	(0.014)	(0.020)	(0.023)	(0.049)	(0.051)	(0.058)
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.8059	0.8059	0.8059	0.8059	0.8059	0.8059	0.8059	0.8059
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.8689	0.9510	0.9379	0.8724	0.8097	0.7698	0.7172	0.6674
	(0.114)	(0.252)	(0.291)	(0.167)	(0.121)	(0.089)	(0.058)	(0.058)
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.8714	0.9641	0.9592	0.8942	0.8494	0.8278	0.8073	0.8059
	(0.112)	(0.244)	(0.279)	(0.153)	(0.101)	(0.051)	(0.008)	(0.0)

TABLE 7

Annualized Certainty Equivalent Returns across Various Portfolios

This table reports the annualized certainty equivalent returns of the implementable combined portfolios and the plug-in portfolios. Entries of the benchmark portfolios are reported to two decimal places when their magnitude exceeds one. The highest certainty equivalent return in each column is shown in bold. N denotes the number of stocks; and $\bar{s}$ denotes the target environmental, social, and governance (ESG) level. The subscripts c-E, c-I, mv, and ef indicate the equality-constrained combined, inequality-constrained combined, plug-in mean-variance, and efficient-frontier portfolios, respectively; a hat and a tilde on a combined portfolio indicate that its combination coefficients are computed with the sample and sophisticated estimators, respectively; and the superscript LW indicates the use of the linear shrinkage covariance matrix of Ledoit and Wolf (2004).

N	6	10	15	25	35	55	75	100
Panel A: $\bar{s} = 0.5630$								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.0586	0.0806	0.0489	-0.0014	0.0183	0.0610	0.0545	0.0494
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.0710	0.1131	0.0796	**0.0601**	**0.0670**	**0.0650**	**0.0613**	**0.0521**
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.1305	0.0908	0.0534	-0.0051	0.0040	0.0537	0.0493	0.0482
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	**0.1517**	**0.1300**	**0.0920**	0.0599	0.0540	0.0574	0.0514	0.0465
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	-0.0913	-0.2373	-0.5562	-1.18	-2.05	-4.78	-8.06	-14.24
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	-92.04	-27.77	-152.68	-1653.91	-46.35	-44.27	-17.62	-58.40
Panel B: $\bar{s} = 0.7018$								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.1189	0.0891	0.0524	-0.0021	0.0133	0.0587	**0.0502**	**0.0482**
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.1393	0.1239	0.0847	**0.0593**	**0.0600**	**0.0609**	0.0501	0.0453
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.1307	0.0931	0.0528	-0.0077	0.0017	0.0521	0.0477	0.0479
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	**0.1525**	**0.1325**	**0.0907**	0.0560	0.0500	0.0548	0.0471	0.0444
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	-0.0188	-0.2313	-0.5539	-1.18	-2.05	-4.78	-8.07	-14.25
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	-19.03	-20.20	-2.57	-70.07	-95.14	-16.38	-11.17	-33.07
Panel C: $\bar{s} = 0.8059$								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.1212	0.0929	0.0530	-0.0040	0.0086	**0.0565**	**0.0468**	**0.0474**
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.1421	0.1288	0.0858	**0.0567**	**0.0529**	0.0562	0.0402	0.0390
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.1255	0.0936	0.0515	-0.0100	-0.0016	0.0514	0.0460	0.0473
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	**0.1470**	**0.1329**	**0.0894**	0.0532	0.0436	0.0529	0.0398	0.0390
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	-0.0128	-0.2300	-0.5548	-1.19	-2.06	-4.79	-8.08	-14.27
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	-1.22	-120.64	-108.93	-266.97	-406.23	-52.08	-56.96	-225.25

TABLE 8

Performance of the ESG-Unaware Portfolios in the Empirical Study

This table reports the average portfolio environmental, social, and governance (ESG) scores (with standard deviations in parentheses) in Panel A and the annualized certainty equivalent return (CER) in Panel B for the ESG-unaware portfolios. In Panel C, for each number of stocks N, the bracketing targets $\bar{s}^-$ and $\bar{s}^+$ are the two of the three evaluated target levels ($\bar{s} \in \{0.5630, 0.7018, 0.8059\}$) closest to the realized average ESG score of $\hat{\boldsymbol{w}}_k^{\mathrm{LW}}$ in Panel A, and each target row is followed by the annualized CERs of $\hat{\boldsymbol{w}}_{c\text{-}E}^{\mathrm{LW}}$ and $\tilde{\boldsymbol{w}}_{c\text{-}E}^{\mathrm{LW}}$ evaluated at that target taken from Table 7. At $N = 35$, the realized score lies marginally below $\bar{s}^-$. The subscripts k and c-E indicate the two-fund portfolio of Kan et al. (2022) and the equality-constrained combined portfolio, respectively; a hat and a tilde on a combined portfolio indicate that its combination coefficients are computed with the sample and sophisticated estimators, respectively; and the superscript LW indicates the use of the linear shrinkage covariance matrix of Ledoit and Wolf (2004), whereas portfolios without it use the sample covariance matrix. Underlined entries are the bracketing target levels.

N	6	10	15	25	35	55	75	100
Panel A: Average portfolio ESG score								
$\hat{\boldsymbol{w}}_k$	0.7778	0.7042	0.6465	0.6400	0.4453	0.5154	0.4818	0.6077
	(0.191)	(0.381)	(0.563)	(0.412)	(0.456)	(0.368)	(0.340)	(0.273)
$\hat{\boldsymbol{w}}_k^{\mathrm{LW}}$	0.7462	0.6739	0.6240	0.6334	0.5588	0.6144	0.6003	0.6291
	(0.147)	(0.226)	(0.351)	(0.241)	(0.222)	(0.177)	(0.130)	(0.071)
Panel B: Annualized certainty equivalent return								
$\hat{\boldsymbol{w}}_k$	0.0913	0.0357	−0.0121	−0.1362	−0.1598	−0.2418	−0.3085	−0.5730
$\hat{\boldsymbol{w}}_k^{\mathrm{LW}}$	0.1113	0.0764	0.0584	−0.0054	0.0120	0.0559	0.0530	0.0482
Panel C: Comparison at the bracketing target levels								
Lower bracketing target $\bar{s}^-$								
$\bar{s}^-$	0.7018	0.5630	0.5630	0.5630	0.5630	0.5630	0.5630	0.5630
$\hat{\boldsymbol{w}}_{c\text{-}E}^{\mathrm{LW}}$	0.1189	0.0806	0.0489	−0.0014	0.0183	0.0610	0.0545	0.0494
$\tilde{\boldsymbol{w}}_{c\text{-}E}^{\mathrm{LW}}$	0.1393	0.1131	0.0796	0.0601	0.0670	0.0650	0.0613	0.0521
Upper bracketing target $\bar{s}^+$								
$\bar{s}^+$	0.8059	0.7018	0.7018	0.7018	0.7018	0.7018	0.7018	0.7018
$\hat{\boldsymbol{w}}_{c\text{-}E}^{\mathrm{LW}}$	0.1212	0.0891	0.0524	−0.0021	0.0133	0.0587	0.0502	0.0482
$\tilde{\boldsymbol{w}}_{c\text{-}E}^{\mathrm{LW}}$	0.1421	0.1239	0.0847	0.0593	0.0600	0.0609	0.0501	0.0453

Appendices

A. Proofs

A.1. Proof of Lemma 1

Proof of Lemma 1. Define the Lagrangian

$$\mathcal{L}(\boldsymbol{w}, \lambda_1, \lambda_2) = \boldsymbol{w}^\top \boldsymbol{\mu} - \frac{\gamma}{2} \boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w} + \lambda_1 (\boldsymbol{w}^\top \boldsymbol{s} - \bar{s}) + \lambda_2 (\boldsymbol{w}^\top \mathbf{1} - 1),$$

where $\lambda_1, \lambda_2 \in \mathbb{R}$ are the Lagrange multipliers. The first-order optimality conditions are:

$$\begin{cases} \dfrac{\partial \mathcal{L}(\boldsymbol{w}, \lambda_1, \lambda_2)}{\partial \boldsymbol{w}} = \boldsymbol{\mu} - \gamma \boldsymbol{\Sigma} \boldsymbol{w} + \lambda_1 \boldsymbol{s} + \lambda_2 \mathbf{1} = \mathbf{0}, \\ \dfrac{\partial \mathcal{L}(\boldsymbol{w}, \lambda_1, \lambda_2)}{\partial \lambda_1} = \boldsymbol{w}^\top \boldsymbol{s} - \bar{s} = 0, \\ \dfrac{\partial \mathcal{L}(\boldsymbol{w}, \lambda_1, \lambda_2)}{\partial \lambda_2} = \boldsymbol{w}^\top \mathbf{1} - 1 = 0, \end{cases}$$

from which we obtain an expression of the optimal portfolio weight:

$$\boldsymbol{w}_{\text{mv}}^* = \frac{1}{\gamma} \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} + \lambda_1 \boldsymbol{s} + \lambda_2 \mathbf{1}).$$

To find the two Lagrange multipliers λ_1 and λ_2, we utilize the two constraints to obtain

$$\begin{cases} \dfrac{1}{\gamma} \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} + \lambda_1 \boldsymbol{s} + \lambda_2 \mathbf{1}) = \bar{s}, \\ \dfrac{1}{\gamma} \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} (\boldsymbol{\mu} + \lambda_1 \boldsymbol{s} + \lambda_2 \mathbf{1}) = 1. \end{cases}$$

When $\mathbf{1}^\top \boldsymbol{\Omega} \boldsymbol{s} \neq 0$ with $\boldsymbol{\Omega} = \boldsymbol{\Sigma}^{-1}(\boldsymbol{s}\mathbf{1}^\top - \mathbf{1}\boldsymbol{s}^\top)\boldsymbol{\Sigma}^{-1}$, we solve the above equations to get

$$\begin{aligned}
\lambda_1 &= \frac{\gamma \mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{s} - \gamma \bar{s}\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1} + \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{s}\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1} - \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}\boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{s}\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{s} - \mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}\boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{s}} \\
&= \frac{\gamma \mathbf{1}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{s} - \bar{s}\mathbf{1}) + \boldsymbol{\mu}^\top \boldsymbol{\Omega}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Omega}\boldsymbol{s}}, \\
\lambda_2 &= \frac{\gamma \bar{s}\boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1} - \gamma \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{s} - \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{s}\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{s} + \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}\boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{s}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{s}\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{s} - \mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}\boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1}\boldsymbol{s}} \\
&= \frac{\gamma \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1}(\bar{s}\mathbf{1} - \boldsymbol{s}) - \boldsymbol{\mu}^\top \boldsymbol{\Omega}\boldsymbol{s}}{\mathbf{1}^\top \boldsymbol{\Omega}\boldsymbol{s}}.
\end{aligned}$$

It is easy to check that $\mathbf{1}^\top \boldsymbol{\Omega}\mathbf{1} = 0$ and $\boldsymbol{s}^\top \boldsymbol{\Omega}\boldsymbol{s} = 0$. We substitute the expressions of λ_1 and λ_2 into the optimal portfolio weight to get

$$\begin{aligned}
\boldsymbol{w}_{\text{mv}}^* &= \frac{1}{\gamma}\boldsymbol{\Sigma}^{-1}\left[\boldsymbol{\mu} + \left(\frac{\gamma \mathbf{1}^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{s} - \bar{s}\mathbf{1}) + \boldsymbol{\mu}^\top \boldsymbol{\Omega}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Omega}\boldsymbol{s}}\right)\boldsymbol{s} + \left(\frac{\gamma \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1}(\bar{s}\mathbf{1} - \boldsymbol{s}) - \boldsymbol{\mu}^\top \boldsymbol{\Omega}\boldsymbol{s}}{\mathbf{1}^\top \boldsymbol{\Omega}\boldsymbol{s}}\right)\mathbf{1}\right] \\
&= \frac{\boldsymbol{\Omega}\boldsymbol{s}}{\mathbf{1}^\top \boldsymbol{\Omega}\boldsymbol{s}} + \bar{s}\frac{-\boldsymbol{\Omega}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Omega}\boldsymbol{s}} + \frac{1}{\gamma}\left(\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - \frac{(\mathbf{1}^\top \boldsymbol{\Omega}\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}\boldsymbol{s} - (\boldsymbol{s}^\top \boldsymbol{\Omega}\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Omega}\boldsymbol{s}}\right) \\
&= \boldsymbol{w}_a + \bar{s}\boldsymbol{w}_e + \frac{1}{\gamma}\boldsymbol{w}_o,
\end{aligned}$$

where

$$\begin{aligned}
\boldsymbol{w}_a &= \frac{\boldsymbol{\Omega}\boldsymbol{s}}{\mathbf{1}^\top \boldsymbol{\Omega}\boldsymbol{s}}, \qquad\qquad \boldsymbol{w}_e = -\frac{\boldsymbol{\Omega}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Omega}\boldsymbol{s}}, \\
\boldsymbol{w}_o &= \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu} - \frac{(\mathbf{1}^\top \boldsymbol{\Omega}\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}\boldsymbol{s} - (\boldsymbol{s}^\top \boldsymbol{\Omega}\boldsymbol{\mu})\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Omega}\boldsymbol{s}}.
\end{aligned}$$

It remains to verify the three properties of the constituent portfolios stated in the lemma.

1. We define the Lagrangian for the optimization problem in part 1 of Lemma 1:

$$\mathcal{L}_1(\boldsymbol{w}, \lambda_1, \lambda_2) = \frac{1}{2}\boldsymbol{w}^\top \boldsymbol{\Sigma}\boldsymbol{w} + \lambda_1 \boldsymbol{w}^\top \boldsymbol{s} + \lambda_2(\boldsymbol{w}^\top \mathbf{1} - 1),$$

where $\lambda_1, \lambda_2 \in \mathbb{R}$ are the Lagrange multipliers. The first-order optimality conditions are

$$\begin{cases} \dfrac{\partial \mathcal{L}_1(\boldsymbol{w}, \lambda_1, \lambda_2)}{\partial \boldsymbol{w}} = \boldsymbol{\Sigma}\boldsymbol{w} + \lambda_1 \boldsymbol{s} + \lambda_2 \mathbf{1} = \mathbf{0}, \\ \dfrac{\partial \mathcal{L}_1(\boldsymbol{w}, \lambda_1, \lambda_2)}{\partial \lambda_1} = \boldsymbol{w}^\top \boldsymbol{s} = 0, \\ \dfrac{\partial \mathcal{L}_1(\boldsymbol{w}, \lambda_1, \lambda_2)}{\partial \lambda_2} = \boldsymbol{w}^\top \mathbf{1} - 1 = 0, \end{cases}$$

from which we obtain an expression for the optimal portfolio weight:

$$\boldsymbol{w}^* = -\boldsymbol{\Sigma}^{-1}(\lambda_1 \boldsymbol{s} + \lambda_2 \mathbf{1}). \tag{27}$$

To find the two Lagrange multipliers, we use the two constraints to get:

$$\begin{cases} \lambda_1 \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s} + \lambda_2 \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1} = 0, \\ \lambda_1 \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s} + \lambda_2 \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1} = -1, \end{cases}$$

which yields

$$\lambda_1 = \frac{-\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s}}{(\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s})^2 - \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1} \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s}},$$
$$\lambda_2 = \frac{\boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s}}{(\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s})^2 - \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1} \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s}}.$$

We strictly substitute the expressions of the two Lagrange multipliers to obtain:

$$\boldsymbol{w}^* = \frac{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s} \boldsymbol{\Sigma}^{-1} \boldsymbol{s} - \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s} \boldsymbol{\Sigma}^{-1} \mathbf{1}}{(\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s})^2 - \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1} \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s}} = \boldsymbol{w}_a.$$

2. We define the Lagrangian for the optimization problem in part 2 of Lemma 1:

$$\mathcal{L}_2(\boldsymbol{w}, \lambda_1, \lambda_2) = \frac{1}{2}\boldsymbol{w}^\top \boldsymbol{\Sigma} \boldsymbol{w} + \lambda_1(\boldsymbol{w}^\top \boldsymbol{s} - 1) + \lambda_2 \boldsymbol{w}^\top \mathbf{1},$$

where $\lambda_1, \lambda_2 \in \mathbb{R}$ are the Lagrange multipliers. As in part 1, the first-order optimality conditions imply that the optimal portfolio weight takes the form in equation (27). The two Lagrange multipliers satisfy

$$\begin{cases} \lambda_1 \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s} + \lambda_2 \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1} = -1, \\ \lambda_1 \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s} + \lambda_2 \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1} = 0, \end{cases}$$

which yields

$$\lambda_1 = \frac{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}}{(\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s})^2 - \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1} \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s}},$$
$$\lambda_2 = \frac{-\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s}}{(\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s})^2 - \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1} \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s}}.$$

Substituting the multipliers leads to the optimal solution:

$$\boldsymbol{w}^* = \frac{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s} \boldsymbol{\Sigma}^{-1} \mathbf{1} - \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1} \boldsymbol{\Sigma}^{-1} \boldsymbol{s}}{(\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s})^2 - \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1} \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s}} = \boldsymbol{w}_e.$$

3. For the zero-weight-factor portfolio $\boldsymbol{w}_o$, it is straightforward to verify:

$$\mathbf{1}^\top \boldsymbol{w}_o = \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \frac{(\mathbf{1}^\top \boldsymbol{\Omega} \boldsymbol{\mu}) \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s} - (\boldsymbol{s}^\top \boldsymbol{\Omega} \boldsymbol{\mu}) \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Omega} \boldsymbol{s}} = 0,$$
$$\boldsymbol{s}^\top \boldsymbol{w}_o = \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{\mu} - \frac{(\mathbf{1}^\top \boldsymbol{\Omega} \boldsymbol{\mu}) \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} \boldsymbol{s} - (\boldsymbol{s}^\top \boldsymbol{\Omega} \boldsymbol{\mu}) \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1} \mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Omega} \boldsymbol{s}} = 0.$$

□

A.2. Proof of Proposition 1

We first establish the following lemma, where $\boldsymbol{A} \circ \boldsymbol{B}$ denotes the Hadamard product (element-wise product) of two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. This lemma is useful in the proof of Proposition 1.

Lemma 2. *For the covariance matrix $\boldsymbol{\Sigma}$, we have the following results:*

1. $\mathbf{1}^\top(\boldsymbol{\Sigma}^{-1} \circ \boldsymbol{\Sigma})\mathbf{1} = N$.
2. $\mathbf{1}^\top \left(\frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}^\top\boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}} \circ \boldsymbol{\Sigma} \right) \mathbf{1} = 1$.

Proof. We denote $\det(\boldsymbol{\Sigma})$ and $\boldsymbol{C}$ as the determinant of $\boldsymbol{\Sigma}$ and the cofactor matrix formed by all of the cofactors of $\boldsymbol{\Sigma}$, respectively. The inverse of $\boldsymbol{\Sigma}$ can be written as:

$$\boldsymbol{\Sigma}^{-1} = \frac{1}{\det(\boldsymbol{\Sigma})}\boldsymbol{C}^\top.$$

Meanwhile, $\det(\boldsymbol{\Sigma})$ can be expressed in terms of its cofactors along any row or column:

$$\det(\boldsymbol{\Sigma}) = \sum_{i=1}^{N} \boldsymbol{\Sigma}_{ij}\boldsymbol{C}_{ij} \text{ along the } j^{\text{th}} \text{ column, or } \sum_{j=1}^{N} \boldsymbol{\Sigma}_{ij}\boldsymbol{C}_{ij} \text{ along the } i^{\text{th}} \text{ row,}$$

where $\boldsymbol{\Sigma}_{ij}$ and $\boldsymbol{C}_{ij}$ denote the $(i,j)^{\text{th}}$ elements of $\boldsymbol{\Sigma}$ and $\boldsymbol{C}$, respectively.

1. The $(i,j)^{\text{th}}$ element in $\boldsymbol{\Sigma}^{-1} \circ \boldsymbol{\Sigma}$ satisfies $(\boldsymbol{\Sigma}^{-1} \circ \boldsymbol{\Sigma})_{ij} = \frac{1}{\det(\boldsymbol{\Sigma})}\boldsymbol{C}_{ji}\boldsymbol{\Sigma}_{ij}$. Therefore, for the symmetric matrix $\boldsymbol{\Sigma}$ with $\boldsymbol{C}^\top = \boldsymbol{C}$, we obtain

$$\mathbf{1}^\top(\boldsymbol{\Sigma}^{-1} \circ \boldsymbol{\Sigma})\mathbf{1} = \sum_{i=1}^{N}\sum_{j=1}^{N} \frac{1}{\det(\boldsymbol{\Sigma})}\boldsymbol{C}_{ji}\boldsymbol{\Sigma}_{ij} = N.$$

2. We denote $u_i = \sum_{k=1}^{N} C_{ik}$ as the sum of the i^{th} row of $\boldsymbol{C}$. Hence, the $(i,j)^{\text{th}}$ element in $\boldsymbol{C}^\top \mathbf{1}\mathbf{1}^\top \boldsymbol{C}^\top$ is obtained by $(\boldsymbol{C}^\top \mathbf{1}\mathbf{1}^\top \boldsymbol{C}^\top)_{ij} = u_i u_j$. Meanwhile, the sum of all elements of $\boldsymbol{C}$ satisfies $\mathbf{1}^\top \boldsymbol{C}\mathbf{1} = \sum_{i=1}^{N} u_i$. Hence, we obtain

$$\begin{aligned}
\mathbf{1}^\top \left(\frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}} \circ \boldsymbol{\Sigma} \right) \mathbf{1} &= \frac{1}{\det(\boldsymbol{\Sigma})} \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{u_i u_j}{\sum_{i=1}^{N} u_i} \boldsymbol{\Sigma}_{ij} \\
&= \frac{1}{\det(\boldsymbol{\Sigma}) \sum_{i=1}^{N} u_i} \sum_{i=1}^{N} u_i \sum_{j=1}^{N} \boldsymbol{\Sigma}_{ij} u_j \\
&= \frac{1}{\det(\boldsymbol{\Sigma}) \sum_{i=1}^{N} u_i} \sum_{i=1}^{N} u_i \det(\boldsymbol{\Sigma}) = 1,
\end{aligned}$$

where the third equality utilizes the fact that $\sum_{j=1}^{N} \boldsymbol{\Sigma}_{ij} u_j = \sum_{k=1}^{N} (\boldsymbol{\Sigma}\boldsymbol{C})_{ik} = \det(\boldsymbol{\Sigma})$. This holds because $\boldsymbol{\Sigma}\boldsymbol{C} = \det(\boldsymbol{\Sigma}) \boldsymbol{I}_N$, where $\boldsymbol{I}_N$ denotes the $N \times N$ identity matrix.

□

Proof of Proposition 1.

1. We first calculate the explicit forms for the expectations $\mathbb{E}[\boldsymbol{\mu}^\top \hat{\boldsymbol{w}}_c]$ and $\mathbb{E}[\boldsymbol{s}^\top \hat{\boldsymbol{w}}_c]$, which are extended from Theorem 1 in Okhrin and Schmid (2006). For ease of reference, we provide detailed calculations below.

   Recall the notation $\hat{\boldsymbol{\Upsilon}} = \hat{\boldsymbol{\Sigma}}^{-1} - \hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}\mathbf{1}^\top \hat{\boldsymbol{\Sigma}}^{-1} / \mathbf{1}^\top \hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}$. The combined portfolio can be written as

$$\hat{\boldsymbol{w}}_c = \frac{\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}}{\mathbf{1}^\top \hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}} + \delta \hat{\boldsymbol{\Upsilon}} \left(\hat{\boldsymbol{\mu}} + \frac{\beta}{\delta} \boldsymbol{s} \right).$$

   Conditional on $\hat{\boldsymbol{\mu}}$, we denote the i^{th} component of the combined portfolio as

$$\hat{w}_i = \hat{w}_{c,i} | \hat{\boldsymbol{\mu}} = \frac{\boldsymbol{e}_i^\top \hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}}{\mathbf{1}^\top \hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}} + \delta \boldsymbol{e}_i^\top \hat{\boldsymbol{\Upsilon}} \left(\hat{\boldsymbol{\mu}} + \frac{\beta}{\delta} \boldsymbol{s} \right),$$

where $\boldsymbol{e}_i$ is an N-dimensional vector with i^{th} element being 1 and other elements 0. Let $\boldsymbol{M} = \left[\boldsymbol{e}_i, \hat{\boldsymbol{\mu}} + \frac{\beta}{\delta}\boldsymbol{s}, \mathbf{1}\right]$ be an $N \times 3$ matrix and $\boldsymbol{M}_1 = \left[\boldsymbol{e}_i, \hat{\boldsymbol{\mu}} + \frac{\beta}{\delta}\boldsymbol{s}\right]$ the first two columns. Denote

$$\hat{\boldsymbol{G}} = \boldsymbol{M}^\top \hat{\boldsymbol{\Sigma}}^{-1} \boldsymbol{M} = \begin{bmatrix} \hat{\boldsymbol{G}}_{11} & \hat{\boldsymbol{G}}_{12} \\ \hat{\boldsymbol{G}}_{21} & \hat{\boldsymbol{G}}_{22} \end{bmatrix} \quad \text{and} \quad \hat{\boldsymbol{H}} = (\hat{\boldsymbol{G}})^{-1} = \begin{bmatrix} \hat{\boldsymbol{H}}_{11} & \hat{\boldsymbol{H}}_{12} \\ \hat{\boldsymbol{H}}_{21} & \hat{\boldsymbol{H}}_{22} \end{bmatrix}, \tag{28}$$

where $\hat{\boldsymbol{G}}_{11}$ is a 2×2 matrix.

We aim to obtain $\hat{w}_i$ using the elements from $\hat{\boldsymbol{H}}$. To achieve this, we utilize the following relationships:

$$\hat{\boldsymbol{G}}_{22} = \mathbf{1}^\top \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{1} = (\hat{\boldsymbol{H}}_{22} - \hat{\boldsymbol{H}}_{21} \hat{\boldsymbol{H}}_{11}^{-1} \hat{\boldsymbol{H}}_{12})^{-1},$$

$$\hat{\boldsymbol{G}}_{12} = -\hat{\boldsymbol{H}}_{11}^{-1} \hat{\boldsymbol{H}}_{12} (\hat{\boldsymbol{H}}_{22} - \hat{\boldsymbol{H}}_{21} \hat{\boldsymbol{H}}_{11}^{-1} \hat{\boldsymbol{H}}_{12})^{-1},$$

$$\hat{\boldsymbol{H}}_{11}^{-1} = \hat{\boldsymbol{G}}_{11} - \hat{\boldsymbol{G}}_{12} \hat{\boldsymbol{G}}_{22}^{-1} \hat{\boldsymbol{G}}_{21} = \boldsymbol{M}_1^\top \hat{\boldsymbol{\Upsilon}} \boldsymbol{M}_1.$$

Firstly, the first element of $-\hat{\boldsymbol{H}}_{11}^{-1} \hat{\boldsymbol{H}}_{12}$ satisfies

$$(-\hat{\boldsymbol{H}}_{11}^{-1} \hat{\boldsymbol{H}}_{12})_1 = \left(\frac{\hat{\boldsymbol{G}}_{12}}{\hat{\boldsymbol{G}}_{22}} \right)_1 = \frac{\boldsymbol{e}_i^\top \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{1}}{\mathbf{1}^\top \hat{\boldsymbol{\Sigma}}^{-1} \mathbf{1}}. \tag{29}$$

Secondly, the first element of $\hat{\boldsymbol{H}}_{11}^{-1} \boldsymbol{\delta}$ where $\boldsymbol{\delta} = [0, \delta]^\top$ satisfies

$$(\hat{\boldsymbol{H}}_{11}^{-1} \boldsymbol{\delta})_1 = \delta \boldsymbol{e}_i^\top \hat{\boldsymbol{\Upsilon}} \left(\hat{\boldsymbol{\mu}} + \frac{\beta}{\delta} \boldsymbol{s} \right). \tag{30}$$

Therefore, $\hat{w}_i$ can be constructed by the first element of $\hat{\boldsymbol{z}} = -\hat{\boldsymbol{H}}_{11}^{-1} (\hat{\boldsymbol{H}}_{12} - \boldsymbol{\delta})$.

Let $\boldsymbol{G}$ and $\boldsymbol{H}$ be the matrices obtained by replacing $\hat{\boldsymbol{\Sigma}}$ with $\boldsymbol{\Sigma}$ in $\hat{\boldsymbol{G}}$ and $\hat{\boldsymbol{H}}$, respectively. Since $T\hat{\boldsymbol{\Sigma}} \sim \mathcal{W}_N(T-1, \boldsymbol{\Sigma})$, Corollary 3.2.6, Theorems 3.2.10 and 3.2.11 of Muirhead

(1982) imply that, conditional on $\hat{\boldsymbol{\mu}}$,

$$T\hat{\boldsymbol{H}} \sim \mathcal{W}_3(T-N+2, \boldsymbol{H}), \tag{31}$$

$$T\hat{\boldsymbol{H}}_{11} \sim \mathcal{W}_2(T-N+2, \boldsymbol{H}_{11}),$$

$$T\hat{\boldsymbol{H}}_{12}|T\hat{\boldsymbol{H}}_{11} \sim \mathcal{N}_2(T\hat{\boldsymbol{H}}_{11}\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}, (\boldsymbol{H}_{22}-\boldsymbol{H}_{21}\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12})T\hat{\boldsymbol{H}}_{11}),$$

where $\mathcal{N}_d$ denotes a d-dimensional normal distribution. Therefore, using the tower property of expectation, we obtain

$$\begin{aligned}\mathbb{E}[\hat{\boldsymbol{z}}] &= \mathbb{E}[\mathbb{E}[(-\hat{\boldsymbol{H}}_{11}^{-1}\hat{\boldsymbol{H}}_{12}+\hat{\boldsymbol{H}}_{11}^{-1}\boldsymbol{\delta})|\hat{\boldsymbol{H}}_{11}]] = \mathbb{E}[-\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}+\hat{\boldsymbol{H}}_{11}^{-1}\boldsymbol{\delta}] \\ &= -\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}+k_1\boldsymbol{H}_{11}^{-1}\boldsymbol{\delta},\end{aligned}$$

where $k_1 = T/(T-N-1)$. Given that the first element of the expectation of $\hat{\boldsymbol{z}}$ is the expectation of $\hat{w}_i$, we have

$$\mathbb{E}[\hat{w}_i] = \frac{\boldsymbol{e}_i^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}} + k_1\delta\boldsymbol{e}_i^\top\boldsymbol{\Upsilon}\left(\hat{\boldsymbol{\mu}}+\frac{\beta}{\delta}\boldsymbol{s}\right). \tag{32}$$

Since $\hat{\boldsymbol{\mu}} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/T)$ and $\hat{\boldsymbol{\mu}}$ is independent of $\hat{\boldsymbol{\Sigma}}$, the unconditional expectation of the portfolio weight is

$$\mathbb{E}[\hat{\boldsymbol{w}}_c] = \mathbb{E}[\mathbb{E}[\hat{\boldsymbol{w}}_c|\hat{\boldsymbol{\mu}}]] = \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}} + k_1\delta\boldsymbol{\Upsilon}\left(\boldsymbol{\mu}+\frac{\beta}{\delta}\boldsymbol{s}\right).$$

Consequently, we obtain the expected OOS portfolio mean and portfolio factor exposure as

follows:

$$\mathbb{E}[\boldsymbol{\mu}^\top \hat{\boldsymbol{w}}_c] = \frac{\boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}} + k_1 \delta \boldsymbol{\mu}^\top \boldsymbol{\Upsilon}\left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s}\right) = \mu_g + k_1\delta\theta_1^2 + k_1\beta\theta_2^2,$$
$$\mathbb{E}[\boldsymbol{s}^\top \hat{\boldsymbol{w}}_c] = \frac{\boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}} + k_1 \delta \boldsymbol{s}^\top \boldsymbol{\Upsilon}\left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s}\right) = \mu_s + k_1\delta\theta_2^2 + k_1\beta\theta_3^2,$$

where $\mu_g = \boldsymbol{\mu}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}/\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}$, $\mu_s = \boldsymbol{s}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}/\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}$, $\theta_1^2 = \boldsymbol{\mu}^\top \boldsymbol{\Upsilon}\boldsymbol{\mu}$, $\theta_2^2 = \boldsymbol{\mu}^\top \boldsymbol{\Upsilon}\boldsymbol{s}$, and $\theta_3^2 = \boldsymbol{s}^\top \boldsymbol{\Upsilon}\boldsymbol{s}$, as defined in the main body of the paper.

2. We then calculate the expectation of the OOS variance $\mathbb{E}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{\Sigma} \hat{\boldsymbol{w}}_c]$. Let σ_{ij} represent the $(i,j)^{\text{th}}$ element of the covariance matrix $\boldsymbol{\Sigma}$. We express the variance as below:

$$\mathbb{E}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{\Sigma} \hat{\boldsymbol{w}}_c] = \sum_{i=1}^{N} \sigma_{ii} \mathbb{E}\left[\mathbb{E}\left[\hat{w}_{c,i}^2 | \hat{\boldsymbol{\mu}}\right]\right] + 2\sum_{i>j} \sigma_{ij} \mathbb{E}\left[\mathbb{E}\left[\hat{w}_{c,i}\hat{w}_{c,j} | \hat{\boldsymbol{\mu}}\right]\right]. \tag{33}$$

We denote $\mathbb{E}\left[\hat{w}_i^2\right] = \mathbb{E}\left[\hat{w}_{c,i}^2 | \hat{\boldsymbol{\mu}}\right]$. Since $\hat{\boldsymbol{z}} = -\hat{\boldsymbol{H}}_{11}^{-1}(\hat{\boldsymbol{H}}_{12} - \boldsymbol{\delta})$ and $\hat{w}_i$ is the first element of $\hat{\boldsymbol{z}}$, it follows from equation (31) that

$$\hat{\boldsymbol{z}} | T\hat{\boldsymbol{H}}_{11} \sim \mathcal{N}_2(-\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12} + \hat{\boldsymbol{H}}_{11}^{-1}\boldsymbol{\delta}, (\boldsymbol{H}_{22} - \boldsymbol{H}_{21}\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12})(T\hat{\boldsymbol{H}}_{11})^{-1}),$$
$$\hat{w}_i | T\hat{\boldsymbol{H}}_{11} \sim \mathcal{N}_1(\mu_{\hat{\boldsymbol{z}},1}, \sigma_{\hat{\boldsymbol{z}},1}^2),$$

where $\mu_{\hat{\boldsymbol{z}},1}$ and $\sigma_{\hat{\boldsymbol{z}},1}^2$ are the first and $(1,1)^{\text{th}}$ elements of the mean vector and covariance matrix of $\hat{\boldsymbol{z}} | T\hat{\boldsymbol{H}}_{11}$, respectively. Therefore,

$$\mathbb{E}[\hat{w}_i^2] = \mathbb{E}[\mathbb{E}[\hat{w}_i^2 | T\hat{\boldsymbol{H}}_{11}]] = \mathbb{E}[\mu_{\hat{\boldsymbol{z}},1}^2] + \mathbb{E}[\sigma_{\hat{\boldsymbol{z}},1}^2]. \tag{34}$$

The first expectation, $\mathbb{E}[\mu^2_{\hat{\boldsymbol{z}},1}]$, is the $(1,1)^{\text{th}}$ element of the following expectation:

$$
\begin{aligned}
&\mathbb{E}[(-\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}+\hat{\boldsymbol{H}}_{11}^{-1}\boldsymbol{\delta})(-\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}+\hat{\boldsymbol{H}}_{11}^{-1}\boldsymbol{\delta})^\top] \\
=&\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{H}_{12}^\top\boldsymbol{H}_{11}^{-1}-\mathbb{E}[\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{\delta}^\top\hat{\boldsymbol{H}}_{11}^{-1}]-\mathbb{E}[\hat{\boldsymbol{H}}_{11}^{-1}\boldsymbol{\delta}\boldsymbol{H}_{12}^\top\boldsymbol{H}_{11}^{-1}]+\mathbb{E}[\hat{\boldsymbol{H}}_{11}^{-1}\boldsymbol{\delta}\boldsymbol{\delta}^\top\hat{\boldsymbol{H}}_{11}^{-1}] \\
=&\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{H}_{12}^\top\boldsymbol{H}_{11}^{-1}-k_1\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12}\boldsymbol{\delta}^\top\boldsymbol{H}_{11}^{-1}-k_1\boldsymbol{H}_{11}^{-1}\boldsymbol{\delta}\boldsymbol{H}_{12}^\top\boldsymbol{H}_{11}^{-1}+Tk_2\boldsymbol{H}_{11}^{-1}\boldsymbol{\delta}\boldsymbol{\delta}^\top\boldsymbol{H}_{11}^{-1} \\
&+k_1k_2\text{tr}(\boldsymbol{H}_{11}^{-1}\boldsymbol{\delta}\boldsymbol{\delta}^\top)\boldsymbol{H}_{11}^{-1},
\end{aligned}
$$

where $k_2 = T/[(T-N)(T-N-3)]$. We apply Theorem 3.3.16 of Gupta and Nagar (2018) when computing the expectations above. Using equations (29) and (30) and the fact that $\text{tr}(\boldsymbol{H}_{11}^{-1}\boldsymbol{\delta}\boldsymbol{\delta}^\top) = \boldsymbol{\delta}^\top\boldsymbol{H}_{11}^{-1}\boldsymbol{\delta}$, we obtain

$$
\begin{aligned}
\mathbb{E}[\mu^2_{\hat{\boldsymbol{z}},1}] = 2k_1\delta\left(\frac{\boldsymbol{e}_i^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}}\right)\left[\boldsymbol{e}_i^\top\boldsymbol{\Upsilon}\left(\hat{\boldsymbol{\mu}}+\frac{\beta}{\delta}\boldsymbol{s}\right)\right]+Tk_2\delta^2\left[\boldsymbol{e}_i^\top\boldsymbol{\Upsilon}\left(\hat{\boldsymbol{\mu}}+\frac{\beta}{\delta}\boldsymbol{s}\right)\right]^2 \\
+\left(\frac{\boldsymbol{e}_i^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}}\right)^2+k_1k_2\delta^2\hat{\psi}^2\boldsymbol{e}_i^\top\boldsymbol{\Upsilon}\boldsymbol{e}_i,
\end{aligned} \tag{35}
$$

where $\hat{\psi}^2 = \left(\hat{\boldsymbol{\mu}}+\frac{\beta}{\delta}\boldsymbol{s}\right)^\top\boldsymbol{\Upsilon}\left(\hat{\boldsymbol{\mu}}+\frac{\beta}{\delta}\boldsymbol{s}\right)$.

The second expectation, $\mathbb{E}[\sigma^2_{\hat{\boldsymbol{z}},1}]$, is the $(1,1)^{\text{th}}$ element of the following expectation:

$$
\mathbb{E}[(\boldsymbol{H}_{22}-\boldsymbol{H}_{21}\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12})(T\hat{\boldsymbol{H}}_{11})^{-1}] = \frac{k_1}{T}(\boldsymbol{H}_{22}-\boldsymbol{H}_{21}\boldsymbol{H}_{11}^{-1}\boldsymbol{H}_{12})\boldsymbol{H}_{11}^{-1},
$$

which means

$$
\mathbb{E}[\sigma^2_{\hat{\boldsymbol{z}},1}] = \frac{k_1}{T}\sigma_g^2\boldsymbol{e}_i^\top\boldsymbol{\Upsilon}\boldsymbol{e}_i, \tag{36}
$$

where $\sigma_g^2 = 1/\mathbf{1}^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}$.

Since $\hat{\boldsymbol{\mu}} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/T)$ and $\hat{\boldsymbol{\mu}}$ is independent of $\hat{\boldsymbol{\Sigma}}$, $\mathbb{E}[\hat{w}_i^2]$ in equation (34) can be calculated by using equations (35) and (36) as follows:

$$
\begin{aligned}
(37) \qquad \mathbb{E}\left[\mathbb{E}\left[\hat{w}_i^2\right]\right] = & \left(\frac{\boldsymbol{e}_i^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}} + k_1\delta\left[\boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s}\right)\right]\right)^2 \\
& + \frac{(T-N+1)}{T} k_1^2 k_2 \delta^2 \left[\boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s}\right)\right]^2 \\
& + \frac{\boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \boldsymbol{e}_i}{T}\left[k_1\sigma_g^2 + k_1 k_2 (T-2)\delta^2 + k_1 k_2 T \delta^2 \psi^2\right],
\end{aligned}
$$

where we apply the expressions of the following two expectations:

$$
\begin{aligned}
(38) \qquad \mathbb{E}\left[\hat{\psi}^2\right] &= \frac{N-1}{T} + \psi^2, \\
\mathbb{E}\left[\left[\boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\left(\hat{\boldsymbol{\mu}} + \frac{\beta}{\delta}\boldsymbol{s}\right)\right]^2\right] &= \frac{\boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \boldsymbol{e}_i}{T} + \left[\boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s}\right)\right]^2.
\end{aligned}
$$

In the above, the first expectation follows from the fact that $T\hat{\psi}^2 \sim \chi^2_{N-1}(T\psi^2)$ with non-centrality parameter $T\psi^2$, where $\psi^2 = \left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s}\right)^\top \boldsymbol{\Upsilon}\left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s}\right)$ and the degrees of freedom $N - 1 = \text{tr}(\boldsymbol{\Upsilon}\boldsymbol{\Sigma}) = \text{rank}(\boldsymbol{\Upsilon}\boldsymbol{\Sigma})$ (Muirhead ((1982), Theorem 1.4.2)). The second expectation follows from the fact that

$\boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\left(\hat{\boldsymbol{\mu}} + \frac{\beta}{\delta}\boldsymbol{s}\right) \sim \mathcal{N}\left(\boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s}\right), \boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\boldsymbol{\Sigma}\boldsymbol{\Upsilon}\boldsymbol{e}_i/T\right)$ (Muirhead ((1982), Theorem 1.2.6)) and $\boldsymbol{\Upsilon}\boldsymbol{\Sigma}\boldsymbol{\Upsilon} = \boldsymbol{\Upsilon}$.

Similarly, we denote $\mathbb{E}[\hat{w}_i\hat{w}_j] = \mathbb{E}\left[\hat{w}_{c,i}\hat{w}_{c,j}|\hat{\boldsymbol{\mu}}\right]$, and let $\boldsymbol{M}^* = \left[\boldsymbol{e}_i, \boldsymbol{e}_j, \hat{\boldsymbol{\mu}} + \frac{\beta}{\delta}\boldsymbol{s}, \mathbf{1}\right]$ be an $N \times 4$ matrix. We partition the matrix $(\boldsymbol{M}^*)^\top \hat{\boldsymbol{\Sigma}}^{-1} \boldsymbol{M}^*$ and its inverse in a manner analogous to equation (28), with the modification that the upper-left block $\hat{\boldsymbol{G}}_{11}^*$ has dimensions 3×3. We adopt similar notations as in the proof of part 1 by adjusting them

with a superscript *. By Corollary 3.2.6 of Muirhead (1982), conditional on $\hat{\boldsymbol{\mu}}$, we have

$$T\hat{\boldsymbol{H}}^* \sim \mathcal{W}_4(T-N+3, \boldsymbol{H}^*),$$

$$T\hat{\boldsymbol{H}}^*_{11} \sim \mathcal{W}_3(T-N+3, \boldsymbol{H}^*_{11}).$$

The portfolio weights $\hat{w}_i$ and $\hat{w}_j$ are the first two elements in $\hat{\boldsymbol{z}}^* = -(\hat{\boldsymbol{H}}^*_{11})^{-1}(\hat{\boldsymbol{H}}^*_{12} - \boldsymbol{\delta}^*)$, where $\boldsymbol{\delta}^* = [0, 0, \delta]^\top$. By the tower property of expectation,

$$\mathbb{E}[\hat{w}_i\hat{w}_j] = \mathbb{E}[\mathbb{E}[\hat{w}_i\hat{w}_j|T\hat{\boldsymbol{H}}^*_{11}]] = \mathbb{E}[\mu_{\hat{\boldsymbol{z}}^*,1}\mu_{\hat{\boldsymbol{z}}^*,2}] + \mathbb{E}[\sigma_{\hat{\boldsymbol{z}}^*,12}], \tag{39}$$

where $\mu_{\hat{\boldsymbol{z}}^*,1}$ and $\mu_{\hat{\boldsymbol{z}}^*,2}$ are the first two elements of the mean vector $\mathbb{E}[\hat{\boldsymbol{z}}^*|T\hat{\boldsymbol{H}}^*_{11}]$ and $\sigma_{\hat{\boldsymbol{z}}^*,12}$ is the $(1,2)^{\text{th}}$ element of the covariance matrix $\mathrm{Cov}[\hat{\boldsymbol{z}}^*|T\hat{\boldsymbol{H}}^*_{11}]$.

Similarly to equations (35) and (36), we obtain

$$\begin{aligned}
\mathbb{E}[\mu_{\hat{\boldsymbol{z}}^*,1}\mu_{\hat{\boldsymbol{z}}^*,2}] &= k_1\delta\left(\frac{\boldsymbol{e}_i^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}}\right)\left[\boldsymbol{e}_j^\top\boldsymbol{\Upsilon}\left(\hat{\boldsymbol{\mu}}+\frac{\beta}{\delta}\boldsymbol{s}\right)\right] + k_1\delta\left(\frac{\boldsymbol{e}_j^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top\boldsymbol{\Sigma}^{-1}\mathbf{1}}\right)\left[\boldsymbol{e}_i^\top\boldsymbol{\Upsilon}\left(\hat{\boldsymbol{\mu}}+\frac{\beta}{\delta}\boldsymbol{s}\right)\right] \\
&\quad + \frac{(\boldsymbol{e}_i^\top\boldsymbol{\Sigma}^{-1}\mathbf{1})(\boldsymbol{e}_j^\top\boldsymbol{\Sigma}^{-1}\mathbf{1})}{(\mathbf{1}^\top\boldsymbol{\Sigma}^{-1}\mathbf{1})^2} + Tk_2\delta^2\left[\boldsymbol{e}_i^\top\boldsymbol{\Upsilon}\left(\hat{\boldsymbol{\mu}}+\frac{\beta}{\delta}\boldsymbol{s}\right)\right]\left[\boldsymbol{e}_j^\top\boldsymbol{\Upsilon}\left(\hat{\boldsymbol{\mu}}+\frac{\beta}{\delta}\boldsymbol{s}\right)\right] + k_1k_2\delta^2\hat{\psi}^2\boldsymbol{e}_i^\top\boldsymbol{\Upsilon}\boldsymbol{e}_j, \\
\mathbb{E}[\sigma_{\hat{\boldsymbol{z}}^*,12}] &= \frac{k_1}{T}\sigma_g^2\boldsymbol{e}_i^\top\boldsymbol{\Upsilon}\boldsymbol{e}_j.
\end{aligned}$$

Since $\hat{\boldsymbol{\mu}} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/T)$ and $\hat{\boldsymbol{\mu}}$ is independent of $\hat{\boldsymbol{\Sigma}}$, we obtain

$$
\begin{aligned}
(40) \qquad & \mathbb{E}\left[\mathbb{E}\left[\hat{w}_i \hat{w}_j\right]\right] \\
&= \left[\frac{\boldsymbol{e}_i^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}} + k_1 \delta \left(\boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s}\right)\right)\right]\left[\frac{\boldsymbol{e}_j^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}} + k_1 \delta \left(\boldsymbol{e}_j^\top \boldsymbol{\Upsilon}\left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s}\right)\right)\right] \\
&\quad + \frac{(T-N+1)}{T} k_1^2 k_2 \delta^2 \left[\boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s}\right)\right]\left[\boldsymbol{e}_j^\top \boldsymbol{\Upsilon}\left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s}\right)\right] \\
&\quad + \frac{\boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \boldsymbol{e}_j}{T}\left[k_1 \sigma_g^2 + k_1 k_2 (T-2)\delta^2 + k_1 k_2 T \delta^2 \psi^2\right],
\end{aligned}
$$

where we utilize the following expectation:

$$
\mathbb{E}\left[\boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\left(\hat{\boldsymbol{\mu}} + \frac{\beta}{\delta}\boldsymbol{s}\right)\left(\hat{\boldsymbol{\mu}} + \frac{\beta}{\delta}\boldsymbol{s}\right)^\top \boldsymbol{\Upsilon} \boldsymbol{e}_j\right] = \boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s}\right)\left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s}\right)^\top \boldsymbol{\Upsilon} \boldsymbol{e}_j + \frac{1}{T}\boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \boldsymbol{e}_j.
$$

Substituting equations (37) and (40) into equation (33), we obtain the expected OOS

portfolio variance as follows:

$$
\begin{aligned}
&\mathbb{E}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{\Sigma} \hat{\boldsymbol{w}}_c] \\
=&\sum_{i=1}^{N} \sigma_{ii} \left[\frac{\boldsymbol{e}_i^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}} + k_1 \delta \boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s} \right) \right]^2 + \sum_{i=1}^{N} \sigma_{ii} \frac{(T-N+1)k_2}{T} \left[k_1 \delta \boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s} \right) \right]^2 \\
&+ 2 \sum_{i>j} \sigma_{ij} \left[\frac{\boldsymbol{e}_i^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}} + k_1 \delta \boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s} \right) \right] \left[\frac{\boldsymbol{e}_j^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}} + k_1 \delta \boldsymbol{e}_j^\top \boldsymbol{\Upsilon} \left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s} \right) \right] \\
&+ 2 \sum_{i>j} \sigma_{ij} \frac{(T-N+1)k_2}{T} \left[k_1 \delta \boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s} \right) \right] \left[k_1 \delta \boldsymbol{e}_j^\top \boldsymbol{\Upsilon} \left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s} \right) \right] \\
&+ \frac{k_1\sigma_g^2 + k_1k_2(T-2)\delta^2 + k_1k_2T\delta^2\psi^2}{T} \left[\sum_{i=1}^{N} \sigma_{ii} \boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \boldsymbol{e}_i + 2 \sum_{i>j} \sigma_{ij} \boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \boldsymbol{e}_j \right] \\
=&\left[\frac{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}} + k_1 \delta \left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s} \right)^\top \boldsymbol{\Upsilon} \right] \boldsymbol{\Sigma} \left[\frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}} + k_1 \delta \boldsymbol{\Upsilon} \left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s} \right) \right] \\
&+ \frac{(T-N+1)k_2}{T} \left[k_1 \delta \left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s} \right)^\top \boldsymbol{\Upsilon} \right] \boldsymbol{\Sigma} \left[k_1 \delta \boldsymbol{\Upsilon} \left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s} \right) \right] \\
&+ \frac{k_1\sigma_g^2 + k_1k_2(T-2)\delta^2 + k_1k_2T\delta^2\psi^2}{T} \mathbf{1}^\top (\boldsymbol{\Upsilon} \circ \boldsymbol{\Sigma}) \mathbf{1} \\
=&\sigma_g^2 + \frac{T + (T-N+1)k_2}{T} k_1^2\delta^2\psi^2 + \frac{N-1}{T}[k_1\sigma_g^2 + k_1k_2(T-2)\delta^2 + k_1k_2T\delta^2\psi^2] \\
=&k_1k_3(\sigma_g^2 + Tk_2\delta^2\psi^2 + k_2(N-1)\delta^2) \\
=&k_1k_3 \left[\sigma_g^2 + Tk_2(\delta^2\theta_1^2 + \beta^2\theta_3^2 + 2\delta\beta\theta_2^2) + k_2(N-1)\delta^2 \right],
\end{aligned}
$$

where $k_3 = (T-2)/T$ and $\boldsymbol{\Upsilon} \circ \boldsymbol{\Sigma}$ represents the Hadamard product (element-wise product) of the two matrices $\boldsymbol{\Upsilon}$ and $\boldsymbol{\Sigma}$. Furthermore, in the above derivation, we apply the following two relationships:

$$
\begin{aligned}
\mathbf{1}^\top \boldsymbol{\Upsilon} \left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s} \right) &= \mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s} \right) - \frac{(\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}) \left[\mathbf{1}^\top \boldsymbol{\Sigma}^{-1} \left(\boldsymbol{\mu} + \frac{\beta}{\delta}\boldsymbol{s} \right) \right]}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}} = 0, \\
\mathbf{1}^\top (\boldsymbol{\Upsilon} \circ \boldsymbol{\Sigma}) \mathbf{1} &= \mathbf{1}^\top \left[\left(\boldsymbol{\Sigma}^{-1} - \frac{\boldsymbol{\Sigma}^{-1}\mathbf{1}\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}} \right) \circ \boldsymbol{\Sigma} \right] \mathbf{1} = N - 1,
\end{aligned}
$$

where the last equation follows from Lemma 2.

□

A.3. Proof of Proposition 2

Proof of Proposition 2. The variance of the OOS portfolio factor exposure can be written as

$$\mathrm{Var}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{s}] = \boldsymbol{s}^\top \mathrm{Var}[\hat{\boldsymbol{w}}_c]\boldsymbol{s}.$$

Hence, it suffices to calculate $\mathrm{Var}[\hat{\boldsymbol{w}}_c]$. We use the notation introduced in Appendix A.2. Let $\hat{w}_i$ denote the i^{th} portfolio weight conditional on $\hat{\boldsymbol{\mu}}$. Using equations (32) and (34), we obtain the conditional variance

$$\begin{aligned}
&\mathrm{Var}[\hat{w}_i] = \mathbb{E}[\hat{w}_i^2] - (\mathbb{E}[\hat{w}_i])^2 \\
=&\left(\frac{\boldsymbol{e}_i^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}}\right)^2 + 2k_1\delta\left(\frac{\boldsymbol{e}_i^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}}\right)\left[\boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\left(\hat{\boldsymbol{\mu}} + \frac{\beta}{\delta}\boldsymbol{s}\right)\right] + Tk_2\delta^2\left[\boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\left(\hat{\boldsymbol{\mu}} + \frac{\beta}{\delta}\boldsymbol{s}\right)\right]^2 \\
&+ \frac{k_1}{T}\sigma_g^2 \boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\boldsymbol{e}_i + k_1k_2\delta^2\hat{\psi}^2\boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\boldsymbol{e}_i - \left[\frac{\boldsymbol{e}_i^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}}{\mathbf{1}^\top \boldsymbol{\Sigma}^{-1}\mathbf{1}} + k_1\delta\boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\left(\hat{\boldsymbol{\mu}} + \frac{\beta}{\delta}\boldsymbol{s}\right)\right]^2 \\
=&\frac{k_1}{T}\sigma_g^2 \boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\boldsymbol{e}_i + k_1k_2\delta^2\hat{\psi}^2\boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\boldsymbol{e}_i + (Tk_2 - k_1^2)\delta^2\left[\boldsymbol{e}_i^\top \boldsymbol{\Upsilon}\left(\hat{\boldsymbol{\mu}} + \frac{\beta}{\delta}\boldsymbol{s}\right)\right]^2.
\end{aligned}$$

Since $\hat{\boldsymbol{\mu}} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}/T)$ and $\hat{\boldsymbol{\mu}}$ is independent of $\hat{\boldsymbol{\Sigma}}$, we utilize equation (38) to obtain

$$
\begin{aligned}
(41) \qquad & \mathrm{Var}[\hat{w}_{c,i}] = \mathbb{E}[\mathrm{Var}[\hat{w}_i]] + \mathrm{Var}[\mathbb{E}[\hat{w}_i]] \\
=& k_1 k_2 \delta^2 \boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \boldsymbol{e}_i \left(\frac{N-1}{T} + \psi^2 \right) + (Tk_2 - k_1^2)\delta^2 \left[\left(\boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \left(\boldsymbol{\mu} + \frac{\beta}{\delta} \boldsymbol{s} \right) \right)^2 + \frac{\boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \boldsymbol{e}_i}{T} \right] \\
& + \frac{k_1}{T} \sigma_g^2 \boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \boldsymbol{e}_i + \frac{k_1^2 \delta^2}{T} \boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \boldsymbol{e}_i \\
=& k_1 k_2 \delta^2 \psi^2 \boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \boldsymbol{e}_i + \left[k_1 \sigma_g^2 + k_1^2 \delta^2 + k_1 k_2 \delta^2 (N-1) + (Tk_2 - k_1^2)\delta^2 \right] \frac{\boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \boldsymbol{e}_i}{T} \\
& + (Tk_2 - k_1^2)\delta^2 \left[\boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \left(\boldsymbol{\mu} + \frac{\beta}{\delta} \boldsymbol{s} \right) \right]^2 .
\end{aligned}
$$

Furthermore, by equations (32) and (39), we obtain

$$
\begin{aligned}
& \mathrm{Cov}[\hat{w}_i, \hat{w}_j] = \mathbb{E}[\hat{w}_i \hat{w}_j] - \mathbb{E}[\hat{w}_i]\mathbb{E}[\hat{w}_j] \\
=& \frac{k_1}{T} \sigma_g^2 \boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \boldsymbol{e}_j + k_1 k_2 \delta^2 \hat{\psi}^2 (\boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \boldsymbol{e}_j) + (Tk_2 - k_1^2)\delta^2 \left[\boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \left(\hat{\boldsymbol{\mu}} + \frac{\beta}{\delta} \boldsymbol{s} \right) \right] \left[\boldsymbol{e}_j^\top \boldsymbol{\Upsilon} \left(\hat{\boldsymbol{\mu}} + \frac{\beta}{\delta} \boldsymbol{s} \right) \right] .
\end{aligned}
$$

By the law of total covariance, we therefore obtain

$$
\begin{aligned}
(42) \qquad & \mathrm{Cov}[\hat{w}_{c,i}, \hat{w}_{c,j}] = \mathbb{E}[\mathrm{Cov}[\hat{w}_i, \hat{w}_j]] + \mathrm{Cov}[\mathbb{E}[\hat{w}_i], \mathbb{E}[\hat{w}_j]] \\
=& k_1 k_2 \delta^2 \psi^2 (\boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \boldsymbol{e}_j) + [k_1 \sigma_g^2 + k_1 k_2 \delta^2 (N-1) + (Tk_2 - k_1^2)\delta^2 + k_1^2 \delta^2] \frac{\boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \boldsymbol{e}_j}{T} \\
& + (Tk_2 - k_1^2)\delta^2 \left[\boldsymbol{e}_i^\top \boldsymbol{\Upsilon} \left(\boldsymbol{\mu} + \frac{\beta}{\delta} \boldsymbol{s} \right) \right] \left[\boldsymbol{e}_j^\top \boldsymbol{\Upsilon} \left(\boldsymbol{\mu} + \frac{\beta}{\delta} \boldsymbol{s} \right) \right] .
\end{aligned}
$$

Combining equations (41) and (42), we obtain the covariance matrix

$$
\mathrm{Var}[\hat{\boldsymbol{w}}_c] = \left[\frac{k_1 \sigma_g^2}{T} + \delta^2 k_1 k_2 (k_3 + \psi^2) \right] \boldsymbol{\Upsilon} + (Tk_2 - k_1^2)\delta^2 \boldsymbol{\Upsilon} \left(\boldsymbol{\mu} + \frac{\beta}{\delta} \boldsymbol{s} \right) \left(\boldsymbol{\mu} + \frac{\beta}{\delta} \boldsymbol{s} \right)^\top \boldsymbol{\Upsilon} .
$$

Consequently, we obtain the variance of the OOS portfolio factor exposure as follows:

$$\mathrm{Var}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{s}] = \boldsymbol{s}^\top \mathrm{Var}[\hat{\boldsymbol{w}}_c]\boldsymbol{s} = \left[\frac{k_1\sigma_g^2}{T} + \delta^2 k_1 k_2 (k_3 + \psi^2)\right]\theta_3^2 + (Tk_2 - k_1^2)\left(\delta\theta_2^2 + \beta\theta_3^2\right)^2,$$

where $\delta^2\psi^2 = \delta^2\theta_1^2 + \beta^2\theta_3^2 + 2\delta\beta\theta_2^2$ and thus, the expression is identical to equation (7). □

A.4. Proof of Proposition 3

Proof of Proposition 3. The proof follows from the standard method of Lagrange multipliers. For the equality-constrained problem in equation (8), the algebraic derivation is identical to the case $\zeta > 0$ of Proposition 6, detailed in Appendix A.9, which yields

$$\delta_E^* = \frac{\theta_1^2\theta_3^2 - \theta_2^4}{\gamma k_2 k_3 \left[T\theta_1^2\theta_3^2 - T\theta_2^4 + \theta_3^2(N-1)\right]},$$
$$\beta_E^* = \frac{-\theta_2^2(\theta_1^2\theta_3^2 - \theta_2^4)}{\gamma k_2 k_3 \theta_3^2 \left[T\theta_1^2\theta_3^2 - T\theta_2^4 + \theta_3^2(N-1)\right]} - \frac{\mu_s - \bar{s}}{k_1\theta_3^2}.$$

These coincide with the expressions stated in the proposition, noting that $D = T(\theta_1^2\theta_3^2 - \theta_2^4) + \theta_3^2(N-1)$. □

A.5. Proof of Proposition 4

Proof of Proposition 4. Define $A = T\theta_1^2 + N - 1$ and $D = T(\theta_1^2\theta_3^2 - \theta_2^4) + \theta_3^2(N-1)$. Then the optimal combination coefficient δ_k^* from Kan et al. (2022), which corresponds to the restriction $\beta = 0$ in equation (4) and is obtained by maximizing the expected OOS MV utility over δ, is given by

$$\delta_k^* = \frac{\theta_1^2}{\gamma k_2 k_3 A}.$$

The corresponding maximum expected OOS utility is

$$\mathbb{E}[U(\hat{\boldsymbol{w}}_k^*)] = \mu_g + k_1\delta_k^*\theta_1^2 - \frac{\gamma}{2}\left[k_1k_3\sigma_g^2 + k_1k_2k_3A(\delta_k^*)^2\right].$$

By Proposition 3, the optimal combination coefficients δ_E^* and β_E^* of the optimal combined three-fund portfolio $\hat{\boldsymbol{w}}_{c\text{-}E}^*$ are

$$\delta_E^* = \frac{\theta_1^2\theta_3^2 - \theta_2^4}{\gamma k_2k_3D} \text{ and } \beta_E^* = \frac{\bar{s} - \mu_s - k_1\theta_2^2\delta_E^*}{k_1\theta_3^2}.$$

Consequently, we use the expressions for moments in Proposition 1 and substitute δ_E^* and β_E^* to obtain the corresponding maximum expected OOS utility as follows:

$$\mathbb{E}[U(\hat{\boldsymbol{w}}_{c\text{-}E}^*)] = \mu_g + k_1\delta_E^*\theta_1^2 + k_1\beta_E^*\theta_2^2 - \frac{\gamma}{2}\left[k_1k_3\sigma_g^2 + k_1k_2k_3A(\delta_E^*)^2 + Tk_1k_2k_3((\beta_E^*)^2\theta_3^2 + 2\delta_E^*\beta_E^*\theta_2^2)\right].$$

Accordingly, we compute the difference

$$\begin{aligned}(43)\quad \Delta_U :=& \mathbb{E}[U(\hat{\boldsymbol{w}}_{c\text{-}E}^*)] - \mathbb{E}[U(\hat{\boldsymbol{w}}_k^*)] \\ =& k_1\theta_1^2(\delta_E^* - \delta_k^*) + k_1\theta_2^2\beta_E^* - \frac{\gamma}{2}k_1k_2k_3\left[A[(\delta_E^*)^2 - (\delta_k^*)^2] + T[\theta_3^2(\beta_E^*)^2 + 2\theta_2^2\delta_E^*\beta_E^*]\right].\end{aligned}$$

We now analyze the terms in the expression for Δ_U given in equation (43). First, since $\delta_k^* = \frac{\theta_1^2}{\gamma k_2k_3A}$, we have $\theta_1^2 = \gamma k_2k_3A\,\delta_k^*$. Substituting this yields:

$$\begin{aligned}(44)\quad k_1\theta_1^2(\delta_E^* - \delta_k^*) - \frac{\gamma}{2}k_1k_2k_3A[(\delta_E^*)^2 - (\delta_k^*)^2] &= \gamma k_1k_2k_3A\left[\delta_k^*(\delta_E^* - \delta_k^*) - \frac{1}{2}[(\delta_E^*)^2 - (\delta_k^*)^2]\right] \\ &= -\frac{\gamma}{2}k_1k_2k_3A(\delta_E^* - \delta_k^*)^2.\end{aligned}$$

Because the coefficient δ_E^* is independent of $\bar{s}$, the above term remains constant with respect to the target $\bar{s}$.

Second, we have $k_1\theta_3^2\beta_E^* = \bar{s} - \mu_s - k_1\theta_2^2\delta_E^*$. Substituting this into the remaining three terms in equation (43) gives

$$
\begin{aligned}
(45) \qquad & k_1\theta_2^2\beta_E^* - \frac{\gamma}{2}k_1k_2k_3T\left[\theta_3^2(\beta_E^*)^2 + 2\theta_2^2\delta_E^*\beta_E^*\right] \\
=& \frac{\theta_2^2(1-\gamma k_2k_3T\delta_E^*)(\bar{s}-\mu_s-k_1\theta_2^2\delta_E^*)}{\theta_3^2} - \frac{\gamma k_2k_3T(\bar{s}-\mu_s-k_1\theta_2^2\delta_E^*)^2}{2k_1\theta_3^2} \\
=& -\frac{\gamma k_2k_3T}{2k_1\theta_3^2}(\bar{s}-\bar{s}^*)^2 + \frac{k_1\theta_2^4\left(1-\gamma k_2k_3T\delta_E^*\right)^2}{2\gamma k_2k_3T\theta_3^2},
\end{aligned}
$$

where we observe that the expression in equation (45) is a concave quadratic function of $\bar{s}$, and its maximum value is attained at

$$
\bar{s}^* = \mu_s + \frac{k_1\theta_2^2}{\gamma k_2k_3T}.
$$

From the analytical forms of δ_E^* and δ_k^*, direct calculation gives

$$
\delta_E^* - \delta_k^* = -\frac{(N-1)\theta_2^4}{\gamma k_2k_3AD}, \qquad 1-\gamma k_2k_3T\delta_E^* = \frac{\theta_3^2(N-1)}{D}.
$$

Substituting these into the constant terms in equations (44) and (45) and using $A\theta_3^2 - T\theta_2^4 = D$, the two constant terms combine to $k_1\theta_2^4(N-1)^2/(2\gamma k_2k_3TAD)$. Finally, combining all these together, we can write the utility difference as follows:

$$
(46) \qquad \Delta_U = -\frac{\gamma k_2k_3T}{2k_1\theta_3^2}\left(\bar{s}-\bar{s}^*\right)^2 + \frac{k_1\theta_2^4(N-1)^2}{2\gamma k_2k_3TAD}.
$$

□

A.6. Proof of Corollary 1

Proof of Corollary 1. By Proposition 4, we have

$$\Delta_U = \Delta_U^{\max} - \frac{\gamma k_2 k_3 T}{2k_1\theta_3^2}(\bar{s} - \bar{s}^*)^2, \text{ where } \Delta_U^{\max} = \frac{k_1\theta_2^4(N-1)^2}{2\gamma k_2 k_3 T A D}. \tag{47}$$

1. For the dominance region $\{\bar{s} : \Delta_U > 0\}$ to be non-empty, the constant term $\Delta_U^{\max}$ in equation (47) must be strictly positive. Since $A > 0$ and $D > 0$ by the Cauchy–Schwarz inequality $\theta_1^2\theta_3^2 \geq \theta_2^4$, this holds if and only if $\theta_2^2 \neq 0$. Conversely, when $\theta_2^2 = 0$, we have $\Delta_U^{\max} = 0$ and $\bar{s}^* = \mu_s$, so that $\Delta_U = -\frac{\gamma k_2 k_3 T}{2k_1\theta_3^2}(\bar{s} - \mu_s)^2 \leq 0$ for all $\bar{s}$, and the dominance region is empty. This proves the first assertion.

2. Suppose $\theta_2^2 \neq 0$. The equation (47) implies that

$$\Delta_U > 0 \quad \Longleftrightarrow \quad (\bar{s} - \bar{s}^*)^2 < \frac{2k_1\theta_3^2}{\gamma k_2 k_3 T}\,\Delta_U^{\max} = \frac{k_1^2\theta_2^4\theta_3^2(N-1)^2}{\gamma^2 k_2^2 k_3^2 T^2 A D}.$$

   Defining the right-hand side as Δ_s^2, we obtain

$$\Delta_U > 0 \quad \Longleftrightarrow \quad |\bar{s} - \bar{s}^*| < \Delta_s,$$

   where Δ_s equals the right-hand side of equation (13). This proves the second assertion.

□

A.7. Calculation of the Sophisticated Estimators

We follow the notation in Appendix A.2, and mark with a dot the quantities defined analogously to those in the proof of Part 1 of Proposition 1. Let $\dot{\boldsymbol{M}} = [\hat{\boldsymbol{\mu}}, \boldsymbol{s}, \mathbf{1}]$ be an $N \times 3$ matrix and $\dot{\boldsymbol{M}}_1 = [\hat{\boldsymbol{\mu}}, \boldsymbol{s}]$ be the first two columns. We partition $\dot{\boldsymbol{M}}^\top \hat{\boldsymbol{\Sigma}}^{-1} \dot{\boldsymbol{M}}$ and its inverse similarly to equation (28), keeping the upper-left block $\dot{\hat{\boldsymbol{G}}}_{11}$ as a 2×2 matrix. We similarly obtain

$$\dot{\hat{\boldsymbol{H}}}_{11}^{-1} = \dot{\hat{\boldsymbol{G}}}_{11} - \dot{\hat{\boldsymbol{G}}}_{12}\dot{\hat{\boldsymbol{G}}}_{22}^{-1}\dot{\hat{\boldsymbol{G}}}_{21} = \dot{\boldsymbol{M}}_1^\top \hat{\boldsymbol{\Upsilon}} \dot{\boldsymbol{M}}_1 = \begin{bmatrix} \hat{\boldsymbol{\mu}}^\top \hat{\boldsymbol{\Upsilon}} \hat{\boldsymbol{\mu}} & \hat{\boldsymbol{\mu}}^\top \hat{\boldsymbol{\Upsilon}} \boldsymbol{s} \\ \boldsymbol{s}^\top \hat{\boldsymbol{\Upsilon}} \hat{\boldsymbol{\mu}} & \boldsymbol{s}^\top \hat{\boldsymbol{\Upsilon}} \boldsymbol{s} \end{bmatrix} = \begin{bmatrix} \hat{\theta}_1^2 & \hat{\theta}_2^2 \\ \hat{\theta}_2^2 & \hat{\theta}_3^2 \end{bmatrix}.$$

Since $T\hat{\boldsymbol{\Sigma}} \sim \mathcal{W}_N(T-1, \boldsymbol{\Sigma})$, Corollary 3.2.6 and Theorem 3.2.11 of Muirhead (1982) imply that, conditional on $\hat{\boldsymbol{\mu}}$,

$$T\dot{\hat{\boldsymbol{H}}} \sim \mathcal{W}_3(T-N+2, \dot{\boldsymbol{H}}),$$

$$T\dot{\hat{\boldsymbol{H}}}_{11} \sim \mathcal{W}_2(T-N+2, \dot{\boldsymbol{H}}_{11}).$$

Therefore, $\mathbb{E}[\dot{\hat{\boldsymbol{H}}}_{11}^{-1} | \hat{\boldsymbol{\mu}}] = k_1 \dot{\boldsymbol{H}}_{11}^{-1}$. Accordingly, we have

$$\mathbb{E}[\hat{\boldsymbol{\mu}}^\top \hat{\boldsymbol{\Upsilon}} \boldsymbol{s}] = \mathbb{E}[\mathbb{E}[\hat{\boldsymbol{\mu}}^\top \hat{\boldsymbol{\Upsilon}} \boldsymbol{s} | \hat{\boldsymbol{\mu}}]] = k_1 \boldsymbol{\mu}^\top \boldsymbol{\Upsilon} \boldsymbol{s},$$

$$\mathbb{E}[\boldsymbol{s}^\top \hat{\boldsymbol{\Upsilon}} \boldsymbol{s}] = k_1 \boldsymbol{s}^\top \boldsymbol{\Upsilon} \boldsymbol{s}.$$

That is, $\mathbb{E}[\hat{\theta}_2^2] = k_1 \theta_2^2$ and $\mathbb{E}[\hat{\theta}_3^2] = k_1 \theta_3^2$ with $k_1 = T/(T-N-1)$. The unbiased estimators of θ_2^2 and θ_3^2 are correspondingly obtained by

$$\tilde{\theta}_2^2 = \frac{T-N-1}{T}\hat{\theta}_2^2 \quad \text{and} \quad \tilde{\theta}_3^2 = \frac{T-N-1}{T}\hat{\theta}_3^2.$$

A.8. Proof of Proposition 5

Proof of Proposition 5.

1. The implementable versions of the optimal combination coefficients δ_E^* and β_E^* in Proposition 3, obtained by replacing each θ_i^2 with its sample estimator $\hat{\theta}_i^2$ $(i = 1, 2, 3)$, are

$$\hat{\delta}_E = \frac{\hat{\theta}_1^2\hat{\theta}_3^2 - \hat{\theta}_2^4}{\gamma k_2 k_3 \left[T\hat{\theta}_1^2\hat{\theta}_3^2 - T\hat{\theta}_2^4 + \hat{\theta}_3^2(N-1)\right]},$$
$$\hat{\beta}_E = \frac{\bar{s} - \hat{\mu}_s - k_1\hat{\theta}_2^2\hat{\delta}_E}{k_1\hat{\theta}_3^2}.$$

Given that $\hat{\theta}_1^2 = \hat{\boldsymbol{\mu}}^\top\hat{\boldsymbol{\Upsilon}}\hat{\boldsymbol{\mu}}$, $\hat{\theta}_2^2 = \hat{\boldsymbol{\mu}}^\top\hat{\boldsymbol{\Upsilon}}\boldsymbol{s}$, $\hat{\theta}_3^2 = \boldsymbol{s}^\top\hat{\boldsymbol{\Upsilon}}\boldsymbol{s}$ and $\hat{\mu}_s = \mathbf{1}^\top\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{s}/\mathbf{1}^\top\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}$, the portfolio factor exposure of the implementable combined portfolio $\hat{\boldsymbol{w}}_{c\text{-}E}$ is calculated as follows:

$$\begin{aligned}\hat{\boldsymbol{w}}_{c\text{-}E}^\top\boldsymbol{s} &= \frac{\mathbf{1}^\top\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{s}}{\mathbf{1}^\top\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}} + \hat{\delta}_E\boldsymbol{s}^\top\hat{\boldsymbol{\Upsilon}}\hat{\boldsymbol{\mu}} + \frac{\bar{s} - \hat{\mu}_s - k_1\hat{\theta}_2^2\hat{\delta}_E}{k_1\hat{\theta}_3^2}\boldsymbol{s}^\top\hat{\boldsymbol{\Upsilon}}\boldsymbol{s} \\ &= \frac{\bar{s} + (k_1 - 1)\hat{\mu}_s}{k_1}.\end{aligned}$$

Since the coefficient of the target factor level $\bar{s}$ in $\hat{\boldsymbol{w}}_{c\text{-}E}^\top\boldsymbol{s}$ is constant (i.e., $1/k_1$), the variance of the portfolio factor exposure $\hat{\boldsymbol{w}}_{c\text{-}E}^\top\boldsymbol{s}$ is independent of $\bar{s}$.

2. The implementable versions of the optimal combination coefficients δ_E^* and β_E^*, obtained by replacing each θ_i^2 with its sophisticated estimator $\tilde{\theta}_i^2$ $(i = 1, 2, 3)$, are

$$\tilde{\delta}_E = \frac{\tilde{\theta}_1^2\tilde{\theta}_3^2 - \tilde{\theta}_2^4}{\gamma k_2 k_3 \left[T\tilde{\theta}_1^2\tilde{\theta}_3^2 - T\tilde{\theta}_2^4 + \tilde{\theta}_3^2(N-1)\right]},$$
$$\tilde{\beta}_E = \frac{\bar{s} - \hat{\mu}_s - k_1\tilde{\theta}_2^2\tilde{\delta}_E}{k_1\tilde{\theta}_3^2}.$$

Since $\tilde{\theta}_2^2 = (T - N - 1)\hat{\theta}_2^2/T$, $\tilde{\theta}_3^2 = (T - N - 1)\hat{\theta}_3^2/T$ and $k_1 = T/(T - N - 1)$, the

portfolio factor exposure of the implementable combined portfolio $\tilde{\boldsymbol{w}}_{c\text{-}E}$ is calculated as follows:

$$\begin{aligned}\tilde{\boldsymbol{w}}_{c\text{-}E}^{\top}\boldsymbol{s} &= \frac{\mathbf{1}^{\top}\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{s}}{\mathbf{1}^{\top}\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{1}} + \tilde{\delta}_E\boldsymbol{s}^{\top}\hat{\boldsymbol{\Upsilon}}\hat{\boldsymbol{\mu}} + \frac{\bar{s} - \hat{\mu}_s - k_1\tilde{\theta}_2^2\tilde{\delta}_E}{k_1\tilde{\theta}_3^2}\boldsymbol{s}^{\top}\hat{\boldsymbol{\Upsilon}}\boldsymbol{s} \\ &= \bar{s},\end{aligned}$$

as desired.

□

A.9. Proof of Proposition 6

Proof of Proposition 6. Let $\boldsymbol{c} = [\delta, \beta]^{\top}$, $\boldsymbol{h} = [k_1\theta_1^2, k_1\theta_2^2]^{\top}$, $O = \mu_g - \frac{\gamma k_1 k_3 \sigma_g^2}{2}$, and $\boldsymbol{A} = \begin{bmatrix} A_1 & A_2 \\ A_2 & A_3 \end{bmatrix}$ with $A_1 = \gamma k_1 k_2 k_3 (T\theta_1^2 + N - 1)$, $A_2 = \gamma T k_1 k_2 k_3 \theta_2^2$ and $A_3 = \gamma T k_1 k_2 k_3 \theta_3^2$. Also, let $\boldsymbol{p} = [k_1\theta_2^2, k_1\theta_3^2]^{\top}$ and $q = \mu_s - \bar{s}$. Then, the objective function of the optimization problem in equation (16) can be written as

$$\begin{aligned}G(\boldsymbol{c}) &= \mu_g - \frac{\gamma k_1 k_3 \sigma_g^2}{2} + k_1\theta_1^2\delta + k_1\theta_2^2\beta - \frac{\gamma k_1 k_2 k_3}{2}[(T\theta_1^2 + N - 1)\delta^2 + T\theta_3^2\beta^2 + 2T\theta_2^2\delta\beta] \\ &= -\frac{1}{2}\boldsymbol{c}^{\top}\boldsymbol{A}\boldsymbol{c} + \boldsymbol{h}^{\top}\boldsymbol{c} + O,\end{aligned}$$

and the constraint becomes

$$\boldsymbol{p}^{\top}\boldsymbol{c} + q \geq 0.$$

Dropping the constant O, which does not affect the optimal solution, we can express the optimization problem as

$$\begin{aligned} \max_{\boldsymbol{c}} \quad & -\frac{1}{2}\boldsymbol{c}^\top \boldsymbol{A}\boldsymbol{c} + \boldsymbol{h}^\top \boldsymbol{c}, \\ \text{s.t.} \quad & \boldsymbol{p}^\top \boldsymbol{c} + q \geq 0. \end{aligned} \tag{48}$$

A symmetric matrix is positive definite if and only if all its eigenvalues are positive. Therefore, since $\boldsymbol{A}$ is a 2×2 symmetric matrix, a necessary and sufficient condition for $\boldsymbol{A}$ to be positive definite is that $A_1 > 0$ and the determinant of $\boldsymbol{A}$ is positive. Given that $N \geq 2$ and $\theta_1^2 = \boldsymbol{\mu}^\top \boldsymbol{\Upsilon} \boldsymbol{\mu}$ is non-negative, we have $A_1 > 0$. Meanwhile, the determinant of $\boldsymbol{A}$ satisfies $\det(\boldsymbol{A}) = \gamma^2 k_1^2 k_2^2 k_3^2 [T^2(\theta_1^2 \theta_3^2 - \theta_2^4) + T(N-1)\theta_3^2] > 0$. Therefore, $\boldsymbol{A}$ is positive definite, and the problem is a convex optimization problem.

The Lagrangian function with a Lagrange multiplier η is

$$\mathcal{L}(\boldsymbol{c}, \eta) = -\frac{1}{2}\boldsymbol{c}^\top \boldsymbol{A}\boldsymbol{c} + \boldsymbol{h}^\top \boldsymbol{c} + \eta(\boldsymbol{p}^\top \boldsymbol{c} + q). \tag{49}$$

The Karush–Kuhn–Tucker (KKT) conditions for this problem are

$$-\boldsymbol{A}\boldsymbol{c} + \boldsymbol{h} + \eta \boldsymbol{p} = \boldsymbol{0}, \tag{50a}$$

$$\boldsymbol{p}^\top \boldsymbol{c} + q \geq 0, \tag{50b}$$

$$\eta \geq 0, \tag{50c}$$

$$\eta(\boldsymbol{p}^\top \boldsymbol{c} + q) = 0. \tag{50d}$$

Since the problem is convex, any point that satisfies the KKT conditions is a global optimum.

Moreover, since $\boldsymbol{A}$ is positive definite, it is invertible, so the condition in equation (50a) implies

$$\boldsymbol{c} = \boldsymbol{A}^{-1}(\boldsymbol{h} + \eta \boldsymbol{p}). \tag{51}$$

We discuss the optimal solution in the following three cases:

- If $\zeta > 0$, we look for a solution with $\eta > 0$. Then, equation (50d) becomes $\boldsymbol{p}^\top \boldsymbol{c} + q = 0$; that is, the inequality constraint is binding, and the optimal solution lies on the boundary. We substitute equation (51) into $\boldsymbol{p}^\top \boldsymbol{c} + q = 0$ to obtain

$$\boldsymbol{p}^\top \boldsymbol{A}^{-1}(\boldsymbol{h} + \eta \boldsymbol{p}) + q = 0,$$

which gives

$$\eta = -\frac{\boldsymbol{p}^\top \boldsymbol{A}^{-1}\boldsymbol{h} + q}{\boldsymbol{p}^\top \boldsymbol{A}^{-1}\boldsymbol{p}} = -\frac{k_1\theta_2^2 + \gamma k_2 k_3 T(\mu_s - \bar{s})}{k_1\theta_3^2} = \zeta > 0, \tag{52}$$

so the dual feasibility condition in equation (50c) holds. The solution is therefore given by

$$\boldsymbol{c} = \boldsymbol{A}^{-1}\left(\boldsymbol{h} + \eta \boldsymbol{p}\right) = \boldsymbol{A}^{-1}\boldsymbol{h} + \eta \boldsymbol{A}^{-1}\boldsymbol{p},$$

which yields the combination coefficients

$$\delta_I^* = \frac{\theta_1^2\theta_3^2 - \theta_2^4}{\gamma k_2 k_3 \left[T\theta_1^2\theta_3^2 - T\theta_2^4 + \theta_3^2(N-1)\right]},$$
$$\beta_I^* = \frac{-\theta_2^2(\theta_1^2\theta_3^2 - \theta_2^4)}{\gamma k_2 k_3 \theta_3^2 \left[T\theta_1^2\theta_3^2 - T\theta_2^4 + \theta_3^2(N-1)\right]} - \frac{\mu_s - \bar{s}}{k_1\theta_3^2}.$$

These expressions are identical to (δ_E^*, β_E^*) in Proposition 3. Moreover, this solution satisfies the primal feasibility condition in equation (50b), since

$\boldsymbol{p}^\top \boldsymbol{c} + q = \boldsymbol{p}^\top \boldsymbol{A}^{-1}(\boldsymbol{h} + \eta \boldsymbol{p}) + q = 0.$

- If $\zeta = 0$, then $k_1\theta_2^2 + \gamma k_2 k_3 T(\mu_s - \bar{s}) = 0$. In this case, we take $\eta = 0$, so equation (50a) implies that the optimal solution is given by

$$\boldsymbol{c} = \boldsymbol{A}^{-1}\boldsymbol{h}. \tag{53}$$

Expanding the above expression, we obtain the combination coefficients

$$\delta_I^* = \frac{\theta_1^2\theta_3^2 - \theta_2^4}{\gamma k_2 k_3 \left[T\theta_1^2\theta_3^2 - T\theta_2^4 + \theta_3^2(N-1)\right]},$$
$$\beta_I^* = \frac{(N-1)\theta_2^2}{\gamma k_2 k_3 T \left[T\theta_1^2\theta_3^2 - T\theta_2^4 + \theta_3^2(N-1)\right]}.$$

This solution satisfies the primal feasibility condition in equation (50b), since

$\boldsymbol{p}^\top \boldsymbol{c} + q = [k_1\theta_2^2 + \gamma k_2 k_3 T(\mu_s - \bar{s})]/(\gamma k_2 k_3 T) = 0.$

Furthermore, because $k_1\theta_2^2 + \gamma k_2 k_3 T(\mu_s - \bar{s}) = 0$ here, the solution in the case $\zeta > 0$ reduces to the one obtained above, so the two cases deliver the same optimal solution. We therefore include this boundary case in the branch $\zeta \geq 0$ of the proposition.

- If $\zeta < 0$, then $[k_1\theta_2^2 + \gamma k_2 k_3 T(\mu_s - \bar{s})] > 0$. In this case, we take the multiplier $\eta = 0$, and the optimal solution is the unconstrained one in equation (53). Unlike the boundary case $\zeta = 0$, the constraint is now strictly slack in the present case, since

$\boldsymbol{p}^\top \boldsymbol{c} + q = [k_1\theta_2^2 + \gamma k_2 k_3 T(\mu_s - \bar{s})]/(\gamma k_2 k_3 T) > 0.$

□

A.10. Proof of Corollary 2

Proof of Corollary 2. Setting $\zeta = -[k_1\theta_2^2 + \gamma k_2 k_3 T(\mu_s - \bar{s})]/(k_1\theta_3^2) = 0$ and solving for $\bar{s}$ yields

$$\bar{s}^* = \mu_s + \frac{k_1\theta_2^2}{\gamma k_2 k_3 T},$$

as defined in equation (12). Since ζ is strictly increasing in $\bar{s}$, $\bar{s} < \bar{s}^*$ corresponds to $\zeta < 0$ and $\bar{s} \geq \bar{s}^*$ to $\zeta \geq 0$. The two assertions then follow from Proposition 6:

- When $\bar{s} < \bar{s}^*$, substituting $\delta_I^* = \delta_E^*$ and $\beta_I^* = (N-1)\theta_2^2/(\gamma k_2 k_3 T D)$ into the expected factor exposure in Proposition 1 gives $\mathbb{E}[(\hat{\boldsymbol{w}}_{c\text{-}I}^*)^\top \boldsymbol{s}] = \mu_s + k_1\theta_2^2/(\gamma k_2 k_3 T) = \bar{s}^* > \bar{s}$.
- When $\bar{s} \geq \bar{s}^*$, the optimal combination coefficients coincide with the equality-constrained ones in Proposition 3, so the expected factor exposure equals $\bar{s}$ by the constraint in equation (8).

□

A.11. Proof of Proposition 7

Proof of Proposition 7. By Proposition 1, the expected OOS MV utility $\mathbb{E}[U(\hat{\boldsymbol{w}}_c)]$ is a quadratic function of the combination coefficients (δ, β), and the expected factor exposure is $\mathbb{E}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{s}] = \mu_s + k_1(\theta_2^2\delta + \theta_3^2\beta)$.

The unconstrained problem coincides with the case $\zeta < 0$ of Proposition 6, where its solution (δ_I^*, β_I^*) is given in closed form. Substituting this solution into the utility $\mathbb{E}[U(\hat{\boldsymbol{w}}_c)]$ from

Proposition 1 and simplifying yields

$$U_{\text{unc}} = \mu_g - \frac{\gamma}{2} k_1 k_3 \sigma_g^2 + \frac{k_1 \left[T\theta_1^2(\theta_1^2\theta_3^2 - \theta_2^4) + (N-1)\theta_2^4 \right]}{2\gamma k_2 k_3 T D}.$$

Substituting (δ_I^*, β_I^*) into the expected factor exposure gives

$\mu_s + k_1(\theta_2^2\delta_I^* + \theta_3^2\beta_I^*) = \mu_s + \frac{k_1\theta_2^2}{\gamma k_2 k_3 T} = \bar{s}^*$.

By Proposition 4, $\mathbb{E}[U(\hat{\boldsymbol{w}}_{c\text{-}E}^*)]$ is a quadratic function of $\bar{s}$ with leading coefficient $-\frac{\gamma k_2 k_3 T}{2k_1\theta_3^2}$. Meanwhile, the constrained utility satisfies $\mathbb{E}[U(\hat{\boldsymbol{w}}_{c\text{-}E}^*)] \leq U_{\text{unc}}$ for every $\bar{s}$, and the bound is attained at $\bar{s} = \bar{s}^*$. Given that a concave quadratic function that attains its maximum value U_{unc} at $\bar{s}^*$ is determined by its leading coefficient, we have

$$\mathbb{E}[U(\hat{\boldsymbol{w}}_{c\text{-}E}^*)] = U_{\text{unc}} - \frac{\gamma k_2 k_3 T}{2k_1\theta_3^2}(\bar{s} - \bar{s}^*)^2,$$

which is equation (19). The two criteria then follow from the equation:

1. Given $U_{\text{unc}} > 0$ and $\alpha \in [0, 1]$, the requirement $\mathbb{E}[U(\hat{\boldsymbol{w}}_{c\text{-}E}^*)] \geq \alpha U_{\text{unc}}$ is equivalent to $\frac{\gamma k_2 k_3 T}{2k_1\theta_3^2}(\bar{s} - \bar{s}^*)^2 \leq (1-\alpha)U_{\text{unc}}$. Solving this inequality for $\bar{s}$ yields the interval in equation (20), and the exact equality condition follows directly from the same algebraic rearrangement.

2. For any $d \geq 0$, the requirement $\mathbb{E}[U(\hat{\boldsymbol{w}}_{c\text{-}E}^*)] \geq U_{\text{unc}} - d$ is equivalent to $\frac{\gamma k_2 k_3 T}{2k_1\theta_3^2}(\bar{s} - \bar{s}^*)^2 \leq d$. Solving this inequality for $\bar{s}$ yields the interval in equation (21), and the exact equality condition follows directly from the same algebraic rearrangement.

□

A.12. Differences between Implementable Portfolios

Let $\hat{\boldsymbol{w}}_{c\text{-}I}$ and $\hat{\boldsymbol{w}}_{c\text{-}E}$ denote the implementable realizations of the optimal combined portfolios $\hat{\boldsymbol{w}}^*_{c\text{-}I}$ and $\hat{\boldsymbol{w}}^*_{c\text{-}E}$, respectively, with their estimated combination coefficients denoted by $(\hat{\delta}_I, \hat{\beta}_I)$ and $(\hat{\delta}_E, \hat{\beta}_E)$. Recall from Proposition 6 that ζ governs the activation of the inequality constraint, and $\hat{\delta}_E = \hat{\delta}_I$ always holds. We now analyze the case where the estimated parameter $\hat{\zeta}$ is strictly negative ($\hat{\zeta} < 0$).

- When $\hat{\zeta} < 0$, $\hat{\boldsymbol{w}}_{c\text{-}I}$ consistently yields a higher portfolio factor exposure than $\hat{\boldsymbol{w}}_{c\text{-}E}$. To verify this, we compute the difference in their factor exposures:

$$
\begin{aligned}
&\hat{\boldsymbol{w}}_{c\text{-}I}^\top \boldsymbol{s} - \hat{\boldsymbol{w}}_{c\text{-}E}^\top \boldsymbol{s} = (\hat{\boldsymbol{w}}_g + \hat{\delta}_I \hat{\boldsymbol{w}}_z + \hat{\beta}_I \hat{\boldsymbol{w}}_s)^\top \boldsymbol{s} - (\hat{\boldsymbol{w}}_g + \hat{\delta}_E \hat{\boldsymbol{w}}_z + \hat{\beta}_E \hat{\boldsymbol{w}}_s)^\top \boldsymbol{s} \\
=&(\hat{\beta}_I - \hat{\beta}_E)\hat{\boldsymbol{w}}_s^\top \boldsymbol{s} = \frac{k_1 \hat{\theta}_2^2 + \gamma k_2 k_3 T(\hat{\mu}_s - \bar{s})}{\gamma k_1 k_2 k_3 T \hat{\theta}_3^2} \hat{\boldsymbol{w}}_s^\top \boldsymbol{s} \\
=&\frac{k_1 \hat{\theta}_2^2 + \gamma k_2 k_3 T(\hat{\mu}_s - \bar{s})}{\gamma k_1 k_2 k_3 T} > 0
\end{aligned}
$$

  where $\hat{\theta}_3^2 = \hat{\boldsymbol{w}}_s^\top \boldsymbol{s} = \boldsymbol{s}^\top \hat{\boldsymbol{\Upsilon}} \boldsymbol{s} > 0$, and the condition $\hat{\zeta} < 0$ implies that $k_1 \hat{\theta}_2^2 + \gamma k_2 k_3 T(\hat{\mu}_s - \bar{s}) > 0$. While the analysis is specific to the sample estimators, the general difference remains $(\hat{\beta}_I - \hat{\beta}_E)\hat{\boldsymbol{w}}_s^\top \boldsymbol{s}$. Because $\hat{\beta}_I - \hat{\beta}_E > 0$ whenever $\hat{\zeta} < 0$, and $\hat{\boldsymbol{w}}_s^\top \boldsymbol{s} > 0$ strictly holds as a quadratic form, this conclusion remains valid when employing sophisticated estimators or shrinkage covariance matrices.

- Next, we evaluate the difference in OOS utility between $\hat{\boldsymbol{w}}_{c\text{-}I}$ and $\hat{\boldsymbol{w}}_{c\text{-}E}$:

$$\begin{aligned}
&\left(\hat{\boldsymbol{w}}_{c\text{-}I}^{\top}\boldsymbol{\mu}-\frac{\gamma}{2}\hat{\boldsymbol{w}}_{c\text{-}I}^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_{c\text{-}I}\right)-\left(\hat{\boldsymbol{w}}_{c\text{-}E}^{\top}\boldsymbol{\mu}-\frac{\gamma}{2}\hat{\boldsymbol{w}}_{c\text{-}E}^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_{c\text{-}E}\right)\\
=&(\hat{\beta}_I-\hat{\beta}_E)\left[\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\mu}-\gamma\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_g-\gamma\hat{\delta}_I\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_z-\frac{\gamma}{2}(\hat{\beta}_I+\hat{\beta}_E)\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_s\right]\\
=&\frac{k_1\hat{\theta}_2^2+\gamma k_2k_3T(\hat{\mu}_s-\bar{s})}{\gamma k_1k_2k_3T\hat{\theta}_3^2}\left[\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\mu}-\gamma\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_{c\text{-}I}+\frac{k_1\hat{\theta}_2^2+\gamma k_2k_3T(\hat{\mu}_s-\bar{s})}{2k_1k_2k_3T\hat{\theta}_3^2}\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_s\right]
\end{aligned}$$

Since $\hat{\zeta}<0$ implies $\frac{k_1\hat{\theta}_2^2+\gamma k_2k_3T(\hat{\mu}_s-\bar{s})}{\gamma k_1k_2k_3T\hat{\theta}_3^2}>0$, the utility difference is positive if and only if the bracketed term is positive:

$$\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\mu}-\gamma\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_{c\text{-}I}+\frac{k_1\hat{\theta}_2^2+\gamma k_2k_3T(\hat{\mu}_s-\bar{s})}{2k_1k_2k_3T\hat{\theta}_3^2}\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_s>0.$$

Rearranging this inequality yields:

$$\gamma\left[(\hat{\mu}_s-\bar{s})(\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_s)-2k_1\hat{\theta}_3^2(\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_{c\text{-}I})\right]>\frac{-k_1}{k_2k_3T}\left[\hat{\theta}_2^2(\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_s)+2k_2k_3T\hat{\theta}_3^2(\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\mu})\right].$$

This leads to two distinct conditional cases for the risk aversion coefficient γ:

  - If $(\hat{\mu}_s-\bar{s})(\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_s)-2k_1\hat{\theta}_3^2(\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_{c\text{-}I})>0$, the OOS utility difference is positive provided that

$$\gamma>\frac{-k_1\left[\hat{\theta}_2^2(\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_s)+2k_2k_3T\hat{\theta}_3^2(\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\mu})\right]}{k_2k_3T\left[(\hat{\mu}_s-\bar{s})(\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_s)-2k_1\hat{\theta}_3^2(\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_{c\text{-}I})\right]}.$$

  - If $(\hat{\mu}_s-\bar{s})(\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_s)-2k_1\hat{\theta}_3^2(\hat{\boldsymbol{w}}_s^{\top}\boldsymbol{\Sigma}\hat{\boldsymbol{w}}_{c\text{-}I})<0$, the OOS utility difference is positive

provided that

$$\gamma < \frac{-k_1 \left[\hat{\theta}_2^2 (\hat{\boldsymbol{w}}_s^\top \boldsymbol{\Sigma} \hat{\boldsymbol{w}}_s) + 2k_2 k_3 T \hat{\theta}_3^2 (\hat{\boldsymbol{w}}_s^\top \boldsymbol{\mu}) \right]}{k_2 k_3 T \left[(\hat{\mu}_s - \bar{s})(\hat{\boldsymbol{w}}_s^\top \boldsymbol{\Sigma} \hat{\boldsymbol{w}}_s) - 2k_1 \hat{\theta}_3^2 (\hat{\boldsymbol{w}}_s^\top \boldsymbol{\Sigma} \hat{\boldsymbol{w}}_{c\text{-}I}) \right]}.$$

Since $\hat{\boldsymbol{w}}_{c\text{-}I}$ depends on γ through its estimated coefficients $\hat{\delta}_I$ and $\hat{\beta}_I$, the two conditions above are implicit in γ rather than explicit thresholds. Hence, the sign of the utility difference depends on the interaction among γ, N, and T.

B. Additional Numerical Results

B.1. List of Stock Tickers

B.1.1. List of the Components from the DJ30

Table B.1 presents the stock tickers for the components from the DJ30 that are considered in the simulation study in Section V.B.

TABLE B.1

List of the Components from the DJ30

This table lists the ticker symbols of the 29 stocks used in the simulation study. They are the constituents of the Dow Jones Industrial Average (DJ30) as of November 8, 2024, excluding Honeywell International Inc. (HON), whose data are unavailable via Wharton Research Data Services.

1	2	3	4	5	6	7	8	9	10
MMM	AMZN	AXP	AMGN	AAPL	BA	CAT	CVX	CSCO	KO
DIS	GS	HD	IBM	JNJ	JPM	MCD	MRK	MSFT	NKE
PG	CRM	TRV	UNH	VZ	V	WMT	NVDA	SHW	

B.1.2. List of the Components from the S&P 500

TABLE B.2

List of the Top 100 Selected Components from the S&P 500

This table lists the ticker symbols of the 100 S&P 500 constituents used in the empirical study. They are the largest constituents by market capitalization as of 2024 among those with complete historical price and environmental, social, and governance (ESG) data and are listed in descending order of market capitalization, from left to right and top to bottom. The asset universe of size N consists of the first N stocks.

1	2	3	4	5	6	7	8	9	10
AAPL	NVDA	MSFT	GOOGL	AMZN	AVGO	BRK-A	WMT	LLY	JPM
V	UNH	XOM	MA	ORCL	COST	HD	PG	NFLX	BAC
JNJ	CRM	CVX	KO	MRK	TMUS	CSCO	WFC	ACN	TMO
AXP	MCD	IBM	MS	DIS	PEP	ABT	ISRG	AMD	GE
GS	PM	ADBE	INTU	QCOM	TXN	DHR	CAT	VZ	BKNG
T	PFE	SPGI	AMGN	BSX	LOW	CMCSA	NEE	SYK	AMAT
PGR	UNP	TJX	C	ETN	SCHW	COP	BA	ADP	FI
BMY	LMT	DE	GILD	MU	NKE	VRTX	ADI	UPS	SBUX
MDT	MMC	CB	PLD	KLAC	SO	EQIX	MO	TT	APH
WM	SHW	ICE	CME	CDNS	INTC	MCO	DUK	AMT	PH

Table B.2 presents the stock tickers for the components from the S&P 500 that are considered in the empirical study in Section V.C.

B.2. Additional Results for the Simulation Study

B.2.1. Additional Results on the Equality and Inequality Formulations

TABLE B.3

The Proportion of Samples Where the Equality- and Inequality-Constrained Rules Differ

This table reports the proportion of the $M = 10{,}000$ simulated samples in which the equality- and inequality-constrained implementable portfolios differ from each other, that is, in which the estimated value of ζ is negative. Each row corresponds to one set of estimators in Table I, and each column to a target factor level $\bar{s}$; the superscript LW denotes the linear shrinkage covariance matrix of Ledoit and Wolf (2004). A smaller proportion indicates that the inequality constraint binds in more samples, in which case the inequality-constrained portfolio coincides with the equality-constrained one.

$\bar{s}$	0.1090	0.1617	0.2889	0.4161	0.4688	0.5734	0.6913	0.8579
$\hat{\boldsymbol{w}}_{c\text{-}E}/\hat{\boldsymbol{w}}_{c\text{-}I}$	0.5685	0.5050	0.3555	0.2179	0.1717	0.1013	0.0507	0.0176
$\tilde{\boldsymbol{w}}_{c\text{-}E}/\tilde{\boldsymbol{w}}_{c\text{-}I}$	0.7753	0.7101	0.5095	0.3124	0.2383	0.1268	0.0512	0.0119
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}/\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.9254	0.8827	0.7181	0.4716	0.3680	0.1919	0.0716	0.0121
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}/\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.9883	0.9778	0.8894	0.6587	0.5219	0.2549	0.0760	0.0064

TABLE B.4

The Proportion of Samples Where the Equality-Constrained Portfolio Attains Higher Utility

This table reports the proportion of simulated samples in which the implementable realization of $\hat{\boldsymbol{w}}^{*}_{c\text{-}E}$ achieves a higher out-of-sample mean-variance utility than that of $\hat{\boldsymbol{w}}^{*}_{c\text{-}I}$, conditional on the two portfolios differing from each other. A larger value indicates that the implementable equality-constrained portfolio outperforms its inequality-constrained counterpart in more samples. Each row corresponds to one set of estimators in Table I, and each column to a target factor level $\bar{s}$; the superscript LW denotes the linear shrinkage covariance matrix of Ledoit and Wolf (2004).

$\bar{s}$	0.1090	0.1617	0.2889	0.4161	0.4688	0.5734	0.6913	0.8579
$\hat{\boldsymbol{w}}_{c\text{-}E}/\hat{\boldsymbol{w}}_{c\text{-}I}$	0.7423	0.8232	0.9463	0.9940	0.9965	1	1	1
$\tilde{\boldsymbol{w}}_{c\text{-}E}/\tilde{\boldsymbol{w}}_{c\text{-}I}$	0.3930	0.5073	0.8057	0.9645	0.9891	1	1	1
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}/\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.7971	0.8647	0.9655	0.9962	0.9981	1	1	1
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}/\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.4810	0.5914	0.8426	0.9727	0.9906	1	1	1

This section presents two tables that support the comparison between the equality- and inequality-constrained implementable portfolios in the simulation study of Section V.B. Table B.3

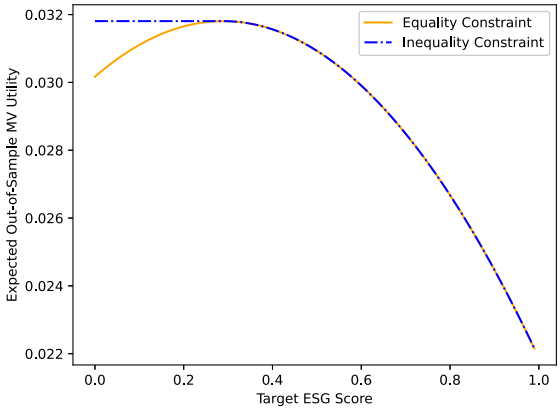
reports the proportion of samples in which the two portfolios differ, and Table B.4 reports, among those samples, the proportion in which the equality-constrained portfolio attains the higher OOS utility.

B.2.2. Expected OOS MV Utilities of the Oracle Portfolios

FIGURE B.1

Expected OOS MV Utilities for the Oracle Combined Portfolios

This figure presents the expected out-of-sample (OOS) mean-variance (MV) utilities of the optimal combined portfolios from the two optimization models: the problem in equation (8) with the equality constraint and the problem in equation (16) with the inequality constraint. The utilities are plotted against different target environmental, social, and governance (ESG) scores ($\bar{s}$). The risk aversion coefficient is set at $\gamma = 5$ and the estimation window is fixed at $T = 120$.



By Proposition 1, the expected OOS MV utility of the combined portfolio $\hat{\boldsymbol{w}}_c$ is

$$\begin{aligned}(54)\qquad \mathbb{E}[U(\hat{\boldsymbol{w}}_c)] =& \mathbb{E}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{\mu}] - \frac{\gamma}{2}\mathbb{E}[\hat{\boldsymbol{w}}_c^\top \boldsymbol{\Sigma} \hat{\boldsymbol{w}}_c] \\ =& \mu_g - \frac{\gamma k_1 k_3 \sigma_g^2}{2} + k_1\theta_1^2\delta + k_1\theta_2^2\beta - \frac{\gamma k_1 k_2 k_3}{2}[(T\theta_1^2 + N - 1)\delta^2 + T\theta_3^2\beta^2 + 2T\theta_2^2\delta\beta].\end{aligned}$$

We replace the combination coefficients δ and β with the optimal values from Proposition 3 (δ_E^*, β_E^*) and Proposition 6 (δ_I^*, β_I^*), which yield the expected OOS MV utilities of $\hat{\boldsymbol{w}}_{c\text{-}E}^*$ and $\hat{\boldsymbol{w}}_{c\text{-}I}^*$,

respectively. Since the population mean vector and covariance matrix are treated as known in the simulation study, the expected OOS utilities can be computed directly for the two oracle combined portfolios. Figure B.1 plots the two utilities against the target ESG score.

B.2.3. Out-of-Sample Mean in the Simulation Study

TABLE B.5

Expected OOS Mean of Portfolio Returns in the Simulation Study

This table reports the expected out-of-sample (OOS) means of portfolio returns for different target factor levels, with the standard deviations in parentheses obtained as the square roots of the expected OOS variances. Panels A–C present the results for the portfolios listed in Table 1. Columns correspond to the target factor levels $\bar{s}$, and the superscript LW denotes the linear shrinkage covariance matrix of Ledoit and Wolf (2004).

$\bar{s}$	0.1090	0.1617	0.2889	0.4161	0.4688	0.5734	0.6913	0.8579
Panel A: Oracle Combined Portfolio								
$\hat{\boldsymbol{w}}^{*}_{c\text{-}E}$	0.0629	0.0621	0.0601	0.0581	0.0573	0.0557	0.0539	0.0513
	(0.1126)	(0.1106)	(0.1064)	(0.1033)	(0.1023)	(0.1010)	(0.1005)	(0.1017)
$\hat{\boldsymbol{w}}^{*}_{c\text{-}I}$	0.0601	0.0601	0.0601	0.0581	0.0573	0.0557	0.0539	0.0513
	(0.1064)	(0.1064)	(0.1064)	(0.1033)	(0.1023)	(0.1010)	(0.1005)	(0.1017)
Panel B: Implementable Combined Portfolio								
Set 1: Apply the Sample Estimators in Equation (14)								
$\hat{\boldsymbol{w}}_{c\text{-}E}$	0.0693	0.0687	0.0672	0.0657	0.0651	0.0639	0.0625	0.0606
	(0.1235)	(0.1224)	(0.1204)	(0.1188)	(0.1184)	(0.1178)	(0.1175)	(0.1181)
$\hat{\boldsymbol{w}}_{c\text{-}I}$	0.0678	0.0675	0.0666	0.0654	0.0649	0.0638	0.0625	0.0606
	(0.1222)	(0.1217)	(0.1203)	(0.1190)	(0.1186)	(0.1179)	(0.1177)	(0.1182)
Set 2: Apply the Sophisticated Estimators in Equation (15)								
$\tilde{\boldsymbol{w}}_{c\text{-}E}$	0.0617	0.0608	0.0589	0.0569	0.0561	0.0545	0.0527	0.0501
	(0.1115)	(0.1095)	(0.1054)	(0.1023)	(0.1013)	(0.1000)	(0.0996)	(0.1007)
$\tilde{\boldsymbol{w}}_{c\text{-}I}$	0.0588	0.0586	0.0577	0.0564	0.0557	0.0543	0.0526	0.0501
	(0.1068)	(0.1062)	(0.1043)	(0.1022)	(0.1014)	(0.1002)	(0.0997)	(0.1008)
Set 3: Apply the Shrinkage Sample Estimators in Equation (22)								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.0536	0.0531	0.0518	0.0506	0.0501	0.0490	0.0478	0.0462
	(0.0934)	(0.0923)	(0.0899)	(0.0882)	(0.0877)	(0.0871)	(0.0870)	(0.0880)
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.0509	0.0508	0.0506	0.0500	0.0497	0.0489	0.0478	0.0462
	(0.0896)	(0.0895)	(0.0889)	(0.0881)	(0.0878)	(0.0872)	(0.0871)	(0.0881)
Set 4: Apply the Shrinkage Sophisticated Estimators in Equation (23)								
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.0452	0.0445	0.0429	0.0412	0.0405	0.0391	0.0376	0.0354
	(0.0839)	(0.0814)	(0.0763)	(0.0724)	(0.0713)	(0.0697)	(0.0693)	(0.0712)
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.0407	0.0407	0.0405	0.0401	0.0398	0.0389	0.0375	0.0354
	(0.0728)	(0.0727)	(0.0724)	(0.0714)	(0.0709)	(0.0698)	(0.0694)	(0.0712)
Panel C: Benchmark Portfolio								
$\hat{\boldsymbol{w}}_{\mathrm{mv}}$	0.1367	0.1359	0.1339	0.1319	0.1311	0.1295	0.1277	0.1251
	(0.2557)	(0.2549)	(0.2531)	(0.2518)	(0.2514)	(0.2509)	(0.2507)	(0.2512)
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	0.1079	0.1072	0.1055	0.1038	0.1031	0.1018	0.1002	0.0980
	(0.1956)	(0.1947)	(0.1928)	(0.1915)	(0.1911)	(0.1907)	(0.1908)	(0.1917)
$\hat{\boldsymbol{w}}_{\mathrm{ef}}$	−0.1890	−0.1721	−0.1312	−0.0904	−0.0734	−0.0399	−0.0020	0.0515
	(79.90)	(72.86)	(55.88)	(38.89)	(31.86)	(17.90)	(2.21)	(20.11)
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	0.0521	0.0472	0.0354	0.0237	0.0188	0.0091	−0.0018	−0.0172
	(28.79)	(26.46)	(20.84)	(15.22)	(12.89)	(8.27)	(3.08)	(4.33)

Table B.5 presents the expected out-of-sample means and standard deviations (reported in

parentheses and obtained as the square roots of the expected out-of-sample variances) in the simulation study.

B.2.4. Results under the Nonlinear Shrinkage Covariance Matrix

The numerical studies in Section V use the linear shrinkage covariance estimator of Ledoit and Wolf (2004). Here, we repeat the analysis using the nonlinear shrinkage estimator of Ledoit and Wolf (2020), with the resulting portfolios denoted with the superscript NL. All other design parameters remain unchanged. Table B.6 parallels Tables 4 and 5 in the main text. While both the linear and nonlinear shrinkage approaches substantially outperform the sample covariance matrix and deliver broadly comparable performance, the linear shrinkage estimator attains a higher expected OOS MV utility at every target level.

TABLE B.6

Performance under the Nonlinear Shrinkage Covariance Matrix in the Simulation Study

This table reports the average portfolio environmental, social, and governance (ESG) scores (with standard deviations in parentheses) and the expected out-of-sample (OOS) mean-variance (MV) utilities of the implementable combined portfolios and the plug-in portfolios constructed with the nonlinear shrinkage covariance matrix of Ledoit and Wolf (2020) (superscript NL). Columns correspond to the target ESG levels $\bar{s}$, and portfolio notation follows Table 1. The plug-in portfolios attain the target exactly in every simulated sample and are therefore omitted from Panel A. The risk aversion coefficient is set at $\gamma = 5$ and the estimation window at $T = 120$.

$\bar{s}$	0.1090	0.1617	0.2889	0.4161	0.4688	0.5734	0.6913	0.8579
Panel A: Average Portfolio ESG Score								
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}E}$	0.2578	0.2973	0.3927	0.4881	0.5276	0.6061	0.6945	0.8195
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.5078	0.5097	0.5213	0.5510	0.5709	0.6236	0.6994	0.8199
	(0.132)	(0.128)	(0.112)	(0.085)	(0.072)	(0.046)	(0.024)	(0.009)
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}E}$	0.1090	0.1617	0.2889	0.4161	0.4688	0.5734	0.6913	0.8579
	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)	(0.0)
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.5056	0.5059	0.5093	0.5268	0.5434	0.6000	0.6966	0.8581
	(0.137)	(0.136)	(0.129)	(0.106)	(0.091)	(0.056)	(0.024)	(0.004)
Panel B: Expected OOS MV Utility								
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}E}$	0.0307	0.0307	0.0305	0.0300	0.0297	0.0290	0.0279	0.0258
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.0296	0.0296	0.0296	0.0294	0.0293	0.0288	0.0278	0.0258
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}E}$	0.0257	0.0260	0.0264	0.0262	0.0259	0.0251	0.0237	0.0209
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.0254	0.0255	0.0255	0.0254	0.0253	0.0248	0.0236	0.0209
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{mv}}$	0.0176	0.0178	0.0180	0.0176	0.0172	0.0162	0.0147	0.0116
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{ef}}$	−4913	−4059	−2333	−1082	−703.9	−194.4	−5.91	−436.5

B.2.5. Utilities under Different Risk Aversion Coefficients

We examine the impact of different risk aversion coefficients ($\gamma = 3$ and 10). The expected OOS utilities are reported in Table B.7. With all other parameters unchanged, we observe that increasing γ generally improves the relative performance of the implementable inequality-constrained portfolio compared with its equality-constrained counterpart. Specifically, when γ is large ($\gamma = 10$), the implementable version of $\hat{\boldsymbol{w}}^*_{c\text{-}I}$ attains a higher utility than that of $\hat{\boldsymbol{w}}^*_{c\text{-}E}$ at target levels where the inequality formulation holds a theoretical advantage ($\bar{s} = 0.1090$ and 0.1617).

B.2.6. Utilities under Different Estimation Windows

We further examine the impact of the estimation window length by considering $T = 60$ and $T = 240$, holding $\gamma = 5$ and the target levels unchanged. The expected OOS utilities are reported in Table B.8. A shorter window amplifies the estimation risk and lowers the utilities of all implementable portfolios; in this case, the shrinkage sample estimators in Set 3 deliver the best performance among the four sets. A longer window reduces estimation risk, raises all utilities, and brings the performance of the four sets closer together.

TABLE B.7

Expected OOS MV Utilities under Different Values of γ

This table reports the expected out-of-sample (OOS) mean-variance (MV) utilities of the implementable combined portfolios in Sets 1–4 of Table 1, together with their counterparts based on the nonlinear shrinkage covariance matrix of Ledoit and Wolf (2020) (superscript NL), under the risk aversion coefficients $\gamma = 3$ and 10. The estimation window is fixed at $T = 120$, and the target factor levels are the same as in Table 5. $\bar{s}$ denotes the target factor level. The benchmark portfolios are defined in Table 1, and the superscript LW denotes the linear shrinkage covariance matrix of Ledoit and Wolf (2004). Entries below "-10^4" are denoted by "—".

$\bar{s}$	0.1090	0.1617	0.2889	0.4161	0.4688	0.5734	0.6913	0.8579
Panel A: $\gamma = 3$								
Set 1: Apply the Sample Estimators in Equation (14)								
$\hat{\boldsymbol{w}}_{c\text{-}E}$	0.0481	0.0479	0.0472	0.0462	0.0458	0.0448	0.0435	0.0414
$\hat{\boldsymbol{w}}_{c\text{-}I}$	0.0473	0.0471	0.0466	0.0458	0.0454	0.0445	0.0433	0.0413
Set 2: Apply the Sophisticated Estimators in Equation (15)								
$\tilde{\boldsymbol{w}}_{c\text{-}E}$	0.0485	0.0484	0.0478	0.0468	0.0462	0.0450	0.0433	0.0404
$\tilde{\boldsymbol{w}}_{c\text{-}I}$	0.0476	0.0475	0.0470	0.0463	0.0458	0.0447	0.0431	0.0403
Set 3: Apply the Shrinkage Sample Estimators in Equation (22)								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.0491	0.0489	0.0483	0.0474	0.0470	0.0461	0.0449	0.0429
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.0480	0.0479	0.0475	0.0469	0.0466	0.0459	0.0448	0.0429
Set 4: Apply the Shrinkage Sophisticated Estimators in Equation (23)								
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.0438	0.0437	0.0432	0.0423	0.0419	0.0408	0.0393	0.0366
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.0424	0.0424	0.0422	0.0417	0.0414	0.0405	0.0392	0.0366
Nonlinear Shrinkage Covariance Matrix, Sample Estimators								
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}E}$	0.0473	0.0471	0.0464	0.0456	0.0452	0.0443	0.0431	0.0412
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.0460	0.0459	0.0456	0.0451	0.0448	0.0441	0.0430	0.0412
Nonlinear Shrinkage Covariance Matrix, Sophisticated Estimators								
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}E}$	0.0406	0.0405	0.0400	0.0392	0.0388	0.0377	0.0362	0.0336
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.0391	0.0391	0.0390	0.0385	0.0382	0.0374	0.0361	0.0336
Benchmark Portfolios								
$\hat{\boldsymbol{w}}_{\mathrm{mv}}$	-0.0471	-0.0472	-0.0479	-0.0488	-0.0494	-0.0506	-0.0523	-0.0552
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	0.0174	0.0172	0.0165	0.0155	0.0149	0.0137	0.0120	0.0091
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{mv}}$	0.0264	0.0262	0.0255	0.0245	0.0240	0.0227	0.0211	0.0182
$\hat{\boldsymbol{w}}_{\mathrm{ef}}$	-9576	-7963	-4684	-2269	-1523	-480.4	-7.324	-606.3
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	-1243	-1050	-651.4	-347.4	-249.3	-102.7	-14.24	-28.17
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{ef}}$	-2948	-2435	-1400	-649.4	-422.3	-116.6	-3.542	-261.9
Panel B: $\gamma = 10$								
Set 1: Apply the Sample Estimators in Equation (14)								
$\hat{\boldsymbol{w}}_{c\text{-}E}$	0.0138	0.0144	0.0155	0.0158	0.0157	0.0153	0.0142	0.0116
$\hat{\boldsymbol{w}}_{c\text{-}I}$	0.0151	0.0152	0.0153	0.0154	0.0154	0.0151	0.0141	0.0116
Set 2: Apply the Sophisticated Estimators in Equation (15)								
$\tilde{\boldsymbol{w}}_{c\text{-}E}$	0.0105	0.0119	0.0144	0.0156	0.0158	0.0154	0.0141	0.0103
$\tilde{\boldsymbol{w}}_{c\text{-}I}$	0.0151	0.0151	0.0152	0.0154	0.0154	0.0152	0.0140	0.0103
Set 3: Apply the Shrinkage Sample Estimators in Equation (22)								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.0145	0.0151	0.0161	0.0164	0.0164	0.0159	0.0149	0.0125
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.0158	0.0158	0.0158	0.0158	0.0158	0.0156	0.0148	0.0125
Set 4: Apply the Shrinkage Sophisticated Estimators in Equation (23)								
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.0093	0.0107	0.0131	0.0144	0.0146	0.0144	0.0132	0.0098
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.0141	0.0141	0.0141	0.0141	0.0141	0.0140	0.0132	0.0098
Nonlinear Shrinkage Covariance Matrix, Sample Estimators								
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}E}$	0.0139	0.0145	0.0155	0.0159	0.0158	0.0154	0.0144	0.0120
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.0153	0.0153	0.0153	0.0153	0.0153	0.0151	0.0143	0.0120
Nonlinear Shrinkage Covariance Matrix, Sophisticated Estimators								
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}E}$	0.0083	0.0097	0.0122	0.0135	0.0137	0.0135	0.0123	0.0090
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.0132	0.0132	0.0132	0.0132	0.0132	0.0131	0.0123	0.0090
Benchmark Portfolios								
$\hat{\boldsymbol{w}}_{\mathrm{mv}}$	-0.0182	-0.0168	-0.0143	-0.0131	-0.0129	-0.0132	-0.0146	-0.0183
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	0.0021	0.0034	0.0056	0.0067	0.0068	0.0065	0.0051	0.0014
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{mv}}$	0.0048	0.0061	0.0083	0.0095	0.0096	0.0092	0.0079	0.0042
$\hat{\boldsymbol{w}}_{\mathrm{ef}}$	—	—	—	-7564	-5075	-1601	-24.41	-2021
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	-4144	-3501	-2171	-1158	-831	-342.2	-47.47	-93.86
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{ef}}$	-9827	-8118	-4666	-2165	-1408	-389	-11.83	-872.8

TABLE B.8

Expected OOS MV Utilities under Different Estimation Windows T

This table reports the expected out-of-sample (OOS) mean-variance (MV) utilities of the implementable combined portfolios in Sets 1–4 of Table 1, together with their counterparts based on the nonlinear shrinkage covariance matrix of Ledoit and Wolf (2020) (superscript NL), under the estimation windows $T = 60$ and $T = 240$. The risk aversion coefficient is fixed at $\gamma = 5$, and the target factor levels $\bar{s}$ are the same as in Table 5. The benchmark portfolios are defined in Table 1, and the superscript LW denotes the linear shrinkage covariance matrix of Ledoit and Wolf (2004).

$\bar{s}$	0.1090	0.1617	0.2889	0.4161	0.4688	0.5734	0.6913	0.8579
Panel A: $T = 60$								
Set 1: Apply the Sample Estimators in Equation (14)								
$\hat{\boldsymbol{w}}_{c\text{-}E}$	0.0152	0.0152	0.0150	0.0146	0.0144	0.0138	0.0130	0.0115
$\hat{\boldsymbol{w}}_{c\text{-}I}$	0.0135	0.0135	0.0135	0.0133	0.0132	0.0128	0.0122	0.0109
Set 2: Apply the Sophisticated Estimators in Equation (15)								
$\tilde{\boldsymbol{w}}_{c\text{-}E}$	0.0139	0.0147	0.0160	0.0163	0.0162	0.0155	0.0140	0.0106
$\tilde{\boldsymbol{w}}_{c\text{-}I}$	0.0146	0.0147	0.0150	0.0151	0.0150	0.0145	0.0134	0.0103
Set 3: Apply the Shrinkage Sample Estimators in Equation (22)								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.0193	0.0192	0.0190	0.0186	0.0184	0.0179	0.0173	0.0160
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.0179	0.0179	0.0179	0.0179	0.0178	0.0176	0.0171	0.0160
Set 4: Apply the Shrinkage Sophisticated Estimators in Equation (23)								
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.0107	0.0113	0.0123	0.0125	0.0124	0.0119	0.0108	0.0082
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.0114	0.0114	0.0114	0.0114	0.0114	0.0113	0.0106	0.0082
Nonlinear Shrinkage Covariance Matrix, Sample Estimators								
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}E}$	0.0179	0.0179	0.0178	0.0174	0.0172	0.0168	0.0161	0.0150
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.0167	0.0167	0.0167	0.0166	0.0166	0.0164	0.0160	0.0149
Nonlinear Shrinkage Covariance Matrix, Sophisticated Estimators								
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}E}$	0.0092	0.0098	0.0109	0.0113	0.0112	0.0108	0.0097	0.0073
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.0102	0.0102	0.0102	0.0102	0.0102	0.0101	0.0096	0.0073
Benchmark Portfolios								
$\hat{\boldsymbol{w}}_{\mathrm{mv}}$	−0.4714	−0.4706	−0.4693	−0.4688	−0.4689	−0.4695	−0.4710	−0.4743
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	−0.0451	−0.0448	−0.0443	−0.0446	−0.0449	−0.0459	−0.0475	−0.0507
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{mv}}$	−0.0284	−0.0280	−0.0275	−0.0277	−0.0280	−0.0289	−0.0304	−0.0336
$\hat{\boldsymbol{w}}_{\mathrm{ef}}$	−527.5	−430.1	−236.2	−100.7	−61.65	−13.83	−7.251	−83.44
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	−750.8	−622.5	−362.6	−173	−115.1	−35.84	−3.593	−61.08
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{ef}}$	−5640	−4689	−2756	−1333	−893.8	−281.3	−5.347	−364.2
Panel B: $T = 240$								
Set 1: Apply the Sample Estimators in Equation (14)								
$\hat{\boldsymbol{w}}_{c\text{-}E}$	0.0418	0.0419	0.0416	0.0410	0.0406	0.0396	0.0382	0.0355
$\hat{\boldsymbol{w}}_{c\text{-}I}$	0.0415	0.0415	0.0414	0.0409	0.0405	0.0396	0.0382	0.0355
Set 2: Apply the Sophisticated Estimators in Equation (15)								
$\tilde{\boldsymbol{w}}_{c\text{-}E}$	0.0415	0.0417	0.0417	0.0411	0.0407	0.0396	0.0380	0.0349
$\tilde{\boldsymbol{w}}_{c\text{-}I}$	0.0414	0.0414	0.0414	0.0409	0.0406	0.0396	0.0380	0.0349
Set 3: Apply the Shrinkage Sample Estimators in Equation (22)								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.0416	0.0416	0.0414	0.0408	0.0404	0.0395	0.0381	0.0355
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.0411	0.0411	0.0410	0.0406	0.0403	0.0394	0.0381	0.0355
Set 4: Apply the Shrinkage Sophisticated Estimators in Equation (23)								
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.0406	0.0408	0.0408	0.0403	0.0399	0.0389	0.0373	0.0343
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.0403	0.0403	0.0403	0.0400	0.0397	0.0388	0.0373	0.0343
Nonlinear Shrinkage Covariance Matrix, Sample Estimators								
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}E}$	0.0409	0.0410	0.0408	0.0402	0.0398	0.0388	0.0374	0.0349
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.0404	0.0404	0.0403	0.0399	0.0396	0.0388	0.0374	0.0349
Nonlinear Shrinkage Covariance Matrix, Sophisticated Estimators								
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}E}$	0.0397	0.0399	0.0399	0.0394	0.0390	0.0380	0.0364	0.0335
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.0393	0.0393	0.0393	0.0391	0.0388	0.0379	0.0364	0.0335
Benchmark Portfolios								
$\hat{\boldsymbol{w}}_{\mathrm{mv}}$	0.0302	0.0304	0.0303	0.0298	0.0294	0.0283	0.0267	0.0236
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	0.0365	0.0366	0.0366	0.0360	0.0356	0.0345	0.0329	0.0298
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{mv}}$	0.0379	0.0380	0.0380	0.0374	0.0370	0.0360	0.0344	0.0313
$\hat{\boldsymbol{w}}_{\mathrm{ef}}$	−13.05	−10.72	−6.05	−2.708	−1.712	−0.4132	−0.0266	−1.428
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	−3112	−2584	−1512	−725.9	−483.7	−148.1	−1.244	−211.9
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{ef}}$	−42.25	−35.1	−20.55	−9.88	−6.591	−2.029	-0.0232	−2.854

B.3. Additional Results for the Empirical Study

B.3.1. Annualized Average Out-of-Sample Returns in the Empirical Study

Table B.9 presents the annualized sample means and standard deviations (reported in parentheses) of OOS portfolio returns in the empirical study.

B.3.2. Results under the Sample and the Nonlinear Shrinkage Covariance Matrices

Tables B.10 and B.11 report the average portfolio ESG scores and the annualized certainty equivalent returns of the portfolios constructed with the sample covariance matrix and with the nonlinear shrinkage covariance estimator of Ledoit and Wolf (2020). They are directly comparable with Tables 6 and 7. Two patterns stand out. First, under the sample covariance matrix, the portfolios based on the sophisticated estimators dominate their sample-form counterparts at every asset universe size, consistent with the motivation of the corrections in Kan and Zhou (2007). Second, for the combined portfolios, the linear shrinkage estimator attains a higher certainty equivalent return than the nonlinear one in every cell, whereas the ordering is reversed for the plug-in portfolio $\hat{\boldsymbol{w}}_{\mathrm{mv}}$ at the three largest asset universe sizes, $N = 55$, 75 and 100.

For the ESG-unaware two-fund portfolio of Kan et al. (2022), the nonlinear version $\hat{\boldsymbol{w}}_k^{\mathrm{NL}}$ attains annualized certainty equivalent returns of 0.1031, 0.0664, 0.0391, -0.0289, 0.0019, 0.0444, 0.0422 and 0.0430 across the eight asset universe sizes, which are lower than those of its linear counterpart in Table 8 at every size.

TABLE B.9

The Annualized Average OOS Returns for Portfolios

This table reports the annualized average out-of-sample (OOS) returns of the portfolios constructed with the linear shrinkage covariance matrix of Ledoit and Wolf (2004) (superscript LW), with the annualized standard deviations in parentheses. Columns correspond to the asset universe size N, Panels A–C to the target environmental, social, and governance (ESG) levels $\bar{s}$, and portfolio notation follows Table 1. The risk aversion coefficient is set at $\gamma = 5$ and the estimation window at $T = 120$.

N	6	10	15	25	35	55	75	100
Panel A: $\bar{s} = 0.5630$								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.3683	0.2352	0.2496	0.1691	0.1932	0.2143	0.1393	0.1034
	(0.3520)	(0.2487)	(0.2834)	(0.2611)	(0.2645)	(0.2476)	(0.1842)	(0.1471)
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.3559	0.1959	0.1557	0.1168	0.1180	0.1149	0.1120	0.1044
	(0.3376)	(0.1820)	(0.1744)	(0.1506)	(0.1428)	(0.1413)	(0.1424)	(0.1446)
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.3280	0.2821	0.2916	0.1759	0.1852	0.2087	0.1347	0.1023
	(0.2811)	(0.2766)	(0.3087)	(0.2691)	(0.2692)	(0.2489)	(0.1848)	(0.1472)
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.3105	0.2437	0.1989	0.1238	0.1096	0.1078	0.1021	0.0987
	(0.2520)	(0.2132)	(0.2068)	(0.1599)	(0.1493)	(0.1421)	(0.1424)	(0.1444)
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	0.4400	0.3392	0.4769	0.3524	0.5103	0.9212	0.6997	0.4717
	(0.4610)	(0.4802)	(0.6428)	(0.7837)	(1.01)	(1.51)	(1.87)	(2.43)
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	-1.46	-1.14	-1.65	-7.20	0.2172	0.1247	-0.7476	-0.9330
	(6.02)	(3.26)	(7.77)	(25.66)	(4.32)	(4.21)	(2.60)	(4.79)
Panel B: $\bar{s} = 0.7018$								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.3298	0.2476	0.2552	0.1679	0.1886	0.2120	0.1351	0.1024
	(0.2904)	(0.2518)	(0.2848)	(0.2608)	(0.2648)	(0.2476)	(0.1843)	(0.1472)
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.3150	0.2096	0.1622	0.1153	0.1113	0.1106	0.1005	0.0977
	(0.2651)	(0.1851)	(0.1761)	(0.1496)	(0.1432)	(0.1409)	(0.1420)	(0.1447)
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.3284	0.2858	0.2912	0.1723	0.1823	0.2070	0.1331	0.1020
	(0.2813)	(0.2776)	(0.3089)	(0.2684)	(0.2688)	(0.2490)	(0.1848)	(0.1472)
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.3110	0.2477	0.1983	0.1195	0.1055	0.1056	0.0979	0.0969
	(0.2518)	(0.2146)	(0.2075)	(0.1594)	(0.1491)	(0.1426)	(0.1426)	(0.1448)
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	0.3991	0.3528	0.4834	0.3509	0.5037	0.9168	0.6882	0.4650
	(0.4088)	(0.4834)	(0.6441)	(0.7835)	(1.01)	(1.51)	(1.87)	(2.43)
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	-0.4494	1.12	0.1368	-1.26	-2.01	-0.1820	1.01	1.16
	(2.73)	(2.92)	(1.04)	(5.25)	(6.10)	(2.55)	(2.21)	(3.70)
Panel C: $\bar{s} = 0.8059$								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.3009	0.2569	0.2594	0.1670	0.1851	0.2102	0.1319	0.1016
	(0.2681)	(0.2562)	(0.2873)	(0.2616)	(0.2657)	(0.2480)	(0.1845)	(0.1473)
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}E}$	0.2843	0.2198	0.1671	0.1142	0.1063	0.1074	0.0919	0.0927
	(0.2385)	(0.1908)	(0.1803)	(0.1516)	(0.1462)	(0.1431)	(0.1438)	(0.1465)
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.3187	0.2889	0.2912	0.1696	0.1789	0.2066	0.1313	0.1015
	(0.2780)	(0.2795)	(0.3096)	(0.2680)	(0.2687)	(0.2491)	(0.1848)	(0.1473)
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{c\text{-}I}$	0.3018	0.2511	0.1988	0.1171	0.1002	0.1049	0.0916	0.0927
	(0.2488)	(0.2174)	(0.2092)	(0.1599)	(0.1505)	(0.1441)	(0.1439)	(0.1465)
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	0.3684	0.3631	0.4883	0.3498	0.4987	0.9136	0.6796	0.4599
	(0.3905)	(0.4871)	(0.6459)	(0.7839)	(1.01)	(1.51)	(1.87)	(2.43)
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	0.3103	2.81	1.47	3.20	-3.67	-0.4120	2.33	2.72
	(0.7829)	(7.03)	(6.65)	(10.40)	(12.69)	(4.55)	(4.87)	(9.55)

TABLE B.10

The Portfolio ESG Scores under the Sample and Nonlinear Shrinkage Covariance Matrices

This table reports the average portfolio environmental, social, and governance (ESG) scores across rolling windows, with standard deviations in parentheses. Portfolios are constructed with the sample covariance matrix or the nonlinear shrinkage covariance matrix of Ledoit and Wolf (2020) (superscript NL). Columns correspond to the asset universe size N, Panels A–C to the target ESG levels $\bar{s}$, and portfolio notation follows Table 1. The equality-constrained portfolios under the sophisticated estimators attain the target $\bar{s}$ exactly in every rolling window, as do the plug-in portfolios; their rows are therefore omitted. The risk aversion coefficient is set at $\gamma = 5$ and the estimation window at $T = 120$.

N	6	10	15	25	35	55	75	100
Panel A: $\bar{s} = 0.5630$								
Sample Covariance Matrix								
$\hat{\boldsymbol{w}}_{c\text{-}E}$	0.5753	0.5616	0.5604	0.5780	0.5737	0.5983	0.6046	0.6512
	(0.005)	(0.011)	(0.021)	(0.038)	(0.042)	(0.087)	(0.102)	(0.185)
$\hat{\boldsymbol{w}}_{c\text{-}I}$	0.8283	0.9674	0.9784	0.9031	0.7873	0.7364	0.6912	0.7276
	(0.250)	(0.520)	(0.587)	(0.459)	(0.357)	(0.266)	(0.231)	(0.210)
$\tilde{\boldsymbol{w}}_{c\text{-}I}$	0.8263	0.9701	0.9847	0.9001	0.7869	0.7326	0.6818	0.7037
	(0.252)	(0.517)	(0.582)	(0.460)	(0.356)	(0.263)	(0.223)	(0.191)
Nonlinear Shrinkage Covariance Matrix								
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}E}$	0.5749	0.5639	0.5641	0.5832	0.5868	0.6148	0.6264	0.6359
	(0.005)	(0.010)	(0.016)	(0.025)	(0.023)	(0.044)	(0.047)	(0.056)
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.8143	0.9044	0.8963	0.8038	0.7099	0.6939	0.6506	0.6425
	(0.229)	(0.436)	(0.457)	(0.288)	(0.190)	(0.132)	(0.073)	(0.062)
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.8125	0.9059	0.9002	0.7973	0.7003	0.6837	0.6371	0.6342
	(0.231)	(0.435)	(0.454)	(0.292)	(0.196)	(0.137)	(0.077)	(0.066)
Panel B: $\bar{s} = 0.7018$								
Sample Covariance Matrix								
$\hat{\boldsymbol{w}}_{c\text{-}E}$	0.7060	0.6877	0.6807	0.6867	0.6709	0.6723	0.6555	0.6732
	(0.005)	(0.011)	(0.021)	(0.038)	(0.042)	(0.087)	(0.102)	(0.185)
$\hat{\boldsymbol{w}}_{c\text{-}I}$	0.8714	1.0256	1.0341	0.9609	0.8507	0.7813	0.7269	0.7397
	(0.210)	(0.476)	(0.548)	(0.418)	(0.318)	(0.239)	(0.215)	(0.206)
$\tilde{\boldsymbol{w}}_{c\text{-}I}$	0.8719	1.0353	1.0495	0.9738	0.8777	0.8183	0.7762	0.7802
	(0.210)	(0.469)	(0.538)	(0.409)	(0.301)	(0.212)	(0.179)	(0.142)
Nonlinear Shrinkage Covariance Matrix								
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}E}$	0.7056	0.6899	0.6844	0.6919	0.6840	0.6888	0.6773	0.6578
	(0.005)	(0.010)	(0.016)	(0.025)	(0.023)	(0.044)	(0.047)	(0.056)
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.8570	0.9647	0.9521	0.8589	0.7677	0.7353	0.6861	0.6586
	(0.188)	(0.391)	(0.418)	(0.247)	(0.153)	(0.102)	(0.057)	(0.057)
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.8577	0.9733	0.9647	0.8674	0.7833	0.7545	0.7136	0.7091
	(0.188)	(0.385)	(0.410)	(0.240)	(0.143)	(0.084)	(0.027)	(0.015)
Panel C: $\bar{s} = 0.8059$								
Sample Covariance Matrix								
$\hat{\boldsymbol{w}}_{c\text{-}E}$	0.8040	0.7823	0.7709	0.7683	0.7438	0.7278	0.6937	0.6897
	(0.005)	(0.011)	(0.021)	(0.038)	(0.042)	(0.087)	(0.102)	(0.185)
$\hat{\boldsymbol{w}}_{c\text{-}I}$	0.9136	1.0747	1.0850	1.0088	0.8996	0.8176	0.7547	0.7492
	(0.181)	(0.441)	(0.516)	(0.387)	(0.290)	(0.220)	(0.204)	(0.203)
$\tilde{\boldsymbol{w}}_{c\text{-}I}$	0.9156	1.0907	1.1097	1.0362	0.9485	0.8907	0.8568	0.8534
	(0.180)	(0.431)	(0.500)	(0.369)	(0.261)	(0.174)	(0.148)	(0.106)
Nonlinear Shrinkage Covariance Matrix								
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}E}$	0.8037	0.7845	0.7746	0.7735	0.7568	0.7443	0.7155	0.6743
	(0.005)	(0.010)	(0.016)	(0.025)	(0.023)	(0.044)	(0.047)	(0.056)
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.8990	1.0161	1.0046	0.9069	0.8165	0.7720	0.7183	0.6743
	(0.159)	(0.355)	(0.385)	(0.215)	(0.126)	(0.082)	(0.050)	(0.056)
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{c\text{-}I}$	0.9012	1.0311	1.0267	0.9298	0.8552	0.8269	0.8062	0.8059
	(0.158)	(0.345)	(0.372)	(0.200)	(0.106)	(0.046)	(0.003)	(0.0)

TABLE B.11

Annualized CERs under the Sample and Nonlinear Shrinkage Covariance Matrices

This table reports the annualized certainty equivalent returns (CERs) of the implementable combined portfolios and the plug-in portfolios constructed with the sample covariance matrix and the nonlinear shrinkage covariance matrix of Ledoit and Wolf (2020) (superscript NL). Columns correspond to the asset universe size N, Panels A–C to the target environmental, social, and governance (ESG) levels $\bar{s}$, and portfolio notation follows Table 1. Entries of the plug-in portfolios are reported to two decimal places when their magnitude exceeds one. Within each panel and for each covariance matrix, the highest certainty equivalent return in each column is shown in bold. The risk aversion coefficient is set at $\gamma = 5$ and the estimation window at $T = 120$.

N	6	10	15	25	35	55	75	100
Panel A: $\bar{s} = 0.5630$								
Sample Covariance Matrix								
$\hat{\boldsymbol{w}}_{c\text{-}E}$	0.0370	0.0544	−0.0091	−0.1130	−0.1355	−0.2040	−0.2787	−0.5383
$\tilde{\boldsymbol{w}}_{c\text{-}E}$	0.0542	**0.1025**	**0.0513**	**0.0373**	**0.0389**	**0.0696**	**0.0075**	−0.2098
$\hat{\boldsymbol{w}}_{c\text{-}I}$	0.1092	0.0425	−0.0336	−0.1416	−0.1711	−0.2324	−0.2992	−0.5555
$\tilde{\boldsymbol{w}}_{c\text{-}I}$	**0.1406**	0.0982	0.0390	0.0172	0.0093	0.0426	−0.0153	**−0.1990**
$\hat{\boldsymbol{w}}_{\text{mv}}$	−0.1284	−0.3224	−0.8112	−2.07	−4.21	−16.37	−53.93	−733.24
$\hat{\boldsymbol{w}}_{\text{ef}}$	−224.99	−0.3265	−2.33	−29.57	−2.48	−39.36	−84.16	−334.02
Nonlinear Shrinkage Covariance Matrix								
$\hat{\boldsymbol{w}}^{\text{NL}}_{c\text{-}E}$	0.0463	0.0747	0.0299	−0.0203	0.0076	0.0499	0.0448	0.0438
$\tilde{\boldsymbol{w}}^{\text{NL}}_{c\text{-}E}$	0.0605	0.1085	0.0637	**0.0490**	**0.0527**	**0.0501**	**0.0522**	**0.0484**
$\hat{\boldsymbol{w}}^{\text{NL}}_{c\text{-}I}$	0.1225	0.0760	0.0284	−0.0294	−0.0052	0.0438	0.0399	0.0428
$\tilde{\boldsymbol{w}}^{\text{NL}}_{c\text{-}I}$	**0.1478**	**0.1158**	**0.0716**	0.0445	0.0413	0.0438	0.0412	0.0431
$\hat{\boldsymbol{w}}^{\text{NL}}_{\text{mv}}$	−0.1089	−0.2471	−0.5959	−1.30	−2.07	−4.28	−6.43	−9.46
$\hat{\boldsymbol{w}}^{\text{NL}}_{\text{ef}}$	−98.22	−16.18	−28.78	−1.49	−2.00	−67.22	−1019.15	−19.03
Panel B: $\bar{s} = 0.7018$								
Sample Covariance Matrix								
$\hat{\boldsymbol{w}}_{c\text{-}E}$	0.1021	0.0628	−0.0056	−0.1142	−0.1398	−0.2038	−0.2820	−0.5373
$\tilde{\boldsymbol{w}}_{c\text{-}E}$	0.1291	**0.1128**	**0.0562**	**0.0361**	**0.0330**	**0.0663**	**−0.0022**	−0.2048
$\hat{\boldsymbol{w}}_{c\text{-}I}$	0.1105	0.0439	−0.0356	−0.1418	−0.1741	−0.2307	−0.3018	−0.5545
$\tilde{\boldsymbol{w}}_{c\text{-}I}$	**0.1426**	0.0999	0.0362	0.0157	0.0034	0.0406	−0.0230	**−0.2042**
$\hat{\boldsymbol{w}}_{\text{mv}}$	−0.0535	−0.3154	−0.8092	−2.08	−4.22	−16.35	−53.93	−733.13
$\hat{\boldsymbol{w}}_{\text{ef}}$	−50.86	−1.50	−25.27	−21.30	−2.66	−123.92	−155.49	−121.38
Nonlinear Shrinkage Covariance Matrix								
$\hat{\boldsymbol{w}}^{\text{NL}}_{c\text{-}E}$	0.1103	0.0831	0.0340	−0.0202	0.0036	**0.0482**	0.0406	**0.0425**
$\tilde{\boldsymbol{w}}^{\text{NL}}_{c\text{-}E}$	0.1332	**0.1187**	**0.0693**	**0.0495**	**0.0473**	0.0470	**0.0411**	0.0408
$\hat{\boldsymbol{w}}^{\text{NL}}_{c\text{-}I}$	0.1232	0.0777	0.0258	−0.0312	−0.0076	0.0430	0.0387	0.0422
$\tilde{\boldsymbol{w}}^{\text{NL}}_{c\text{-}I}$	**0.1491**	0.1178	0.0683	0.0420	0.0378	0.0418	0.0382	0.0399
$\hat{\boldsymbol{w}}^{\text{NL}}_{\text{mv}}$	−0.0331	−0.2401	−0.5929	−1.30	−2.08	−4.28	−6.44	−9.47
$\hat{\boldsymbol{w}}^{\text{NL}}_{\text{ef}}$	−21.98	−65.79	−4.91	−6.38	−32.63	−19.12	−21.67	−59.35
Panel C: $\bar{s} = 0.8059$								
Sample Covariance Matrix								
$\hat{\boldsymbol{w}}_{c\text{-}E}$	0.1082	0.0665	−0.0050	−0.1164	−0.1441	−0.2042	−0.2849	−0.5367
$\tilde{\boldsymbol{w}}_{c\text{-}E}$	0.1371	**0.1173**	**0.0570**	**0.0330**	**0.0263**	**0.0615**	**−0.0123**	**−0.2055**
$\hat{\boldsymbol{w}}_{c\text{-}I}$	0.1069	0.0434	−0.0385	−0.1428	−0.1779	−0.2295	−0.3034	−0.5537
$\tilde{\boldsymbol{w}}_{c\text{-}I}$	**0.1388**	0.0991	0.0324	0.0132	−0.0031	0.0386	−0.0295	−0.2076
$\hat{\boldsymbol{w}}_{\text{mv}}$	−0.0454	−0.3134	−0.8104	−2.08	−4.22	−16.33	−53.93	−733.05
$\hat{\boldsymbol{w}}_{\text{ef}}$	−1.50	−4.03	−79.28	−73.83	−11.15	−422.24	−231.42	−145.68
Nonlinear Shrinkage Covariance Matrix								
$\hat{\boldsymbol{w}}^{\text{NL}}_{c\text{-}E}$	0.1153	0.0867	0.0350	−0.0213	−0.0003	**0.0465**	**0.0372**	**0.0415**
$\tilde{\boldsymbol{w}}^{\text{NL}}_{c\text{-}E}$	0.1392	**0.1231**	**0.0707**	**0.0478**	**0.0413**	0.0430	0.0312	0.0340
$\hat{\boldsymbol{w}}^{\text{NL}}_{c\text{-}I}$	0.1184	0.0774	0.0234	−0.0325	−0.0101	0.0420	0.0366	**0.0415**
$\tilde{\boldsymbol{w}}^{\text{NL}}_{c\text{-}I}$	**0.1442**	0.1171	0.0650	0.0395	0.0323	0.0398	0.0311	0.0340
$\hat{\boldsymbol{w}}^{\text{NL}}_{\text{mv}}$	−0.0248	−0.2381	−0.5934	−1.30	−2.09	−4.28	−6.45	−9.48
$\hat{\boldsymbol{w}}^{\text{NL}}_{\text{ef}}$	−0.5905	−137.93	−52.14	−20.28	−105.55	−106.45	−262.65	−203.90

B.3.3. Utilities under Different Risk Aversion Coefficients

We repeat the empirical analysis with the risk aversion coefficients $\gamma = 3$ and 10, holding the estimation window $T = 120$ and the target levels unchanged. Table B.12 reports the annualized certainty equivalent returns under both risk aversion coefficients. Varying the risk aversion coefficient shifts the level of the certainty equivalent returns but leaves the comparisons across estimators and formulations in Section V.C essentially unchanged.

TABLE B.12

Annualized Certainty Equivalent Returns under Alternative Risk Aversion Coefficients

This table reports the annualized certainty equivalent returns under the risk aversion coefficients $\gamma = 3$ and 10, with the estimation window fixed at $T = 120$ and the target level fixed at $\bar{s} = 0.7018$; the patterns at the other two target levels $(0.5630, 0.8059)$ are the same. The baseline case $\gamma = 5$ is reported in Tables 7 and B.11. Entries of the plug-in portfolios are reported to two decimal places when their magnitude exceeds one. Within each panel, the highest certainty equivalent return in each column is shown in bold. The superscripts LW and NL denote the linear and nonlinear shrinkage covariance matrices of Ledoit and Wolf (2004) and Ledoit and Wolf (2020), respectively. Columns correspond to the asset universe size N, and portfolio notation follows Table 1.

N	6	10	15	25	35	55	75	100
Panel A: $\gamma = 3$								
Linear Shrinkage Covariance Matrix								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{c},E}$	0.1824	0.1060	0.0635	−0.0219	0.0043	0.0824	0.0706	**0.0705**
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{c},E}$	0.2141	0.1593	0.1138	**0.0777**	**0.0838**	**0.0857**	**0.0707**	0.0663
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{c},I}$	0.1902	0.1165	0.0676	−0.0293	−0.0150	0.0721	0.0668	0.0699
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{c},I}$	**0.2237**	**0.1778**	**0.1272**	0.0742	0.0670	0.0761	0.0667	0.0654
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	−0.0333	−0.4158	−0.9353	−1.98	−3.44	−7.99	−13.48	−23.76
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	−11.60	−11.67	−1.49	−42.55	−57.89	−9.90	−6.30	−19.38
Nonlinear Shrinkage Covariance Matrix								
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{c},E}$	0.1718	0.0990	0.0397	−0.0460	−0.0037	0.0731	0.0603	0.0622
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{c},E}$	0.2087	0.1547	0.0961	0.0682	0.0710	0.0702	0.0607	0.0601
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{c},I}$	0.1807	0.0937	0.0309	−0.0620	−0.0221	0.0651	0.0574	0.0618
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{c},I}$	0.2217	0.1569	0.0999	0.0578	0.0553	0.0627	0.0572	0.0592
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{mv}}$	−0.0567	−0.4307	−0.9963	−2.18	−3.48	−7.13	−10.75	−15.78
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{ef}}$	−12.82	−40.14	−2.75	−3.49	−19.59	−10.95	−12.58	−36.17
Panel B: $\gamma = 10$								
Linear Shrinkage Covariance Matrix								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{c},E}$	−0.0325	0.0349	0.0043	−0.0225	−0.0145	0.0071	**−0.0002**	**−0.0050**
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{c},E}$	−0.0232	0.0556	0.0229	**0.0103**	**0.0077**	**0.0084**	−0.0005	−0.0070
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{c},I}$	0.0040	0.0339	0.0026	−0.0265	−0.0206	0.0033	−0.0018	−0.0052
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{c},I}$	**0.0174**	**0.0569**	**0.0242**	0.0077	0.0020	0.0044	−0.0031	−0.0082
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	−0.1143	−0.1347	−0.3079	−0.6213	−1.05	−2.41	−4.05	−7.16
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	−37.61	−41.51	−5.27	−138.88	−188.28	−32.57	−23.35	−67.30
Nonlinear Shrinkage Covariance Matrix								
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{c},E}$	−0.0400	0.0293	−0.0106	−0.0365	−0.0253	−0.0036	−0.0075	−0.0058
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{c},E}$	−0.0303	0.0496	0.0085	−0.0003	−0.0051	−0.0038	−0.0070	−0.0074
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{c},I}$	−0.0023	0.0234	−0.0175	−0.0430	−0.0308	−0.0069	−0.0088	−0.0061
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{c},I}$	0.0129	0.0458	0.0056	−0.0052	−0.0101	−0.0073	−0.0099	−0.0085
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{mv}}$	−0.1223	−0.1392	−0.3311	−0.6822	−1.06	−2.17	−3.24	−4.77
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{ef}}$	−44.89	−129.90	−10.32	−13.60	−65.25	−39.53	−44.39	−117.30

B.3.4. Utilities under Different Estimation Windows

We repeat the empirical analysis with the estimation windows $T = 60$ and $T = 240$, holding $\gamma = 5$ and the target levels unchanged. Each estimation window consists of the most recent T weeks before the corresponding investment week. For $T = 60$, the estimation window for the first investment week of 2014 is formed from the historical returns prior to this date, which extend back to 2011-09-19, in the same way as in the baseline case $T = 120$. The OOS weeks are therefore identical to those of the baseline, and the results are comparable cell by cell.

For $T = 60$, the asset universe sizes are restricted to $N \leq 35$. The moment conditions underlying our expressions require $T > N + 3$, which already rules out $N \geq 75$. At $N = 55$ this condition is formally satisfied, but the ratio N/T is close to one, and the estimation error then dominates the combination coefficients and the resulting portfolios become unstable, so we omit this case as well. For $T = 240$, the historical returns provide only 120 weeks before the first investment week of 2014, so the estimation window cannot be formed there. The first investment week is instead deferred to the first week preceded by 240 observations, and the OOS period starts on 2016-04-25 and is accordingly shorter than that of the baseline.

Table B.13 reports the annualized certainty equivalent returns under both estimation windows. The estimation window governs the size of the sophisticated corrections, which grow with N relative to T. Under the shorter window $T = 60$, the corrections are large, and the sample-form estimators generally outperform the sophisticated estimators at $N = 35$, compared with $N = 75$ in the baseline case $T = 120$. Under the longer window $T = 240$, the corrections are mild, and the sophisticated estimators deliver higher certainty equivalent returns in nearly

every case. Both patterns are consistent with the double adjustment discussed in the second observation of Section V.C.

TABLE B.13

Annualized Certainty Equivalent Returns under Alternative Estimation Windows

This table reports the annualized certainty equivalent returns under the estimation windows $T = 60$ and $T = 240$, with the risk aversion coefficient fixed at $\gamma = 5$ and the target level fixed at $\bar{s} = 0.7018$; the patterns at the other two target levels $(0.5630, 0.8059)$ are the same. The baseline case $T = 120$ is reported in Tables 7 and B.11. For $T = 60$, the out-of-sample (OOS) weeks are identical to the baseline, and the entries for $N \geq 55$ are omitted (denoted by "—"). For $T = 240$, the OOS period starts in April 2016. Entries of the plug-in portfolios are reported to two decimal places when their magnitude exceeds one. Within each panel, the highest certainty equivalent return in each column is shown in bold. The superscripts LW and NL denote the linear and nonlinear shrinkage covariance matrices of Ledoit and Wolf (2004) and Ledoit and Wolf (2020), respectively. Columns correspond to the asset universe size N, and portfolio notation follows Table 1.

N	6	10	15	25	35	55	75	100
Panel A: $T = 60$								
Linear Shrinkage Covariance Matrix								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\text{c-}E}$	0.1007	0.0411	0.0729	0.0616	**0.0627**	—	—	—
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{\text{c-}E}$	**0.1344**	0.1035	0.1016	**0.0786**	0.0593	—	—	—
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\text{c-}I}$	0.0866	0.0490	0.0763	0.0489	0.0571	—	—	—
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{\text{c-}I}$	0.1245	**0.1187**	**0.1107**	0.0686	0.0550	—	—	—
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	−0.2144	−0.7038	−1.33	−2.87	−5.22	—	—	—
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	−4.57	−191.56	−657.23	−54.47	−436.15	—	—	—
Nonlinear Shrinkage Covariance Matrix								
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\text{c-}E}$	0.0681	0.0175	0.0517	0.0340	0.0552	—	—	—
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{\text{c-}E}$	0.1085	0.0862	0.0874	0.0642	0.0549	—	—	—
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\text{c-}I}$	0.0418	0.0180	0.0546	0.0172	0.0508	—	—	—
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{\text{c-}I}$	0.0888	0.0944	0.1006	0.0510	0.0515	—	—	—
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{mv}}$	−0.2937	−0.8039	−1.52	−3.34	−5.46	—	—	—
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{ef}}$	−8.91	−309.65	−14282.43	−69.50	−1.55	—	—	—
Panel B: $T = 240$								
Linear Shrinkage Covariance Matrix								
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\text{c-}E}$	0.1047	0.1249	0.0980	0.0521	0.0327	−0.0144	0.0052	0.0026
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{\text{c-}E}$	0.1021	**0.1280**	**0.1192**	**0.0781**	**0.0659**	**0.0370**	**0.0562**	**0.0377**
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\text{c-}I}$	0.1061	0.1153	0.0905	0.0438	0.0174	−0.0241	0.0034	0.0016
$\tilde{\boldsymbol{w}}^{\mathrm{LW}}_{\text{c-}I}$	**0.1073**	0.1213	0.1145	0.0715	0.0501	0.0286	0.0548	0.0371
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{mv}}$	0.0641	0.0154	−0.1672	−0.4422	−0.8474	−1.83	−2.99	−5.04
$\hat{\boldsymbol{w}}^{\mathrm{LW}}_{\mathrm{ef}}$	−1359.73	−24.35	−1.66	−0.4861	−0.6522	−2.19	−8.52	−42.13
Nonlinear Shrinkage Covariance Matrix								
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\text{c-}E}$	0.0999	0.1197	0.0854	0.0367	0.0222	−0.0134	0.0127	0.0117
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{\text{c-}E}$	0.0957	0.1241	0.1087	0.0705	0.0570	0.0331	0.0480	0.0307
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\text{c-}I}$	0.0949	0.1026	0.0697	0.0241	0.0074	−0.0218	0.0115	0.0116
$\tilde{\boldsymbol{w}}^{\mathrm{NL}}_{\text{c-}I}$	0.0958	0.1103	0.0976	0.0606	0.0423	0.0258	0.0470	0.0306
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{mv}}$	0.0609	0.0054	−0.1870	−0.4953	−0.8740	−1.77	−2.69	−4.11
$\hat{\boldsymbol{w}}^{\mathrm{NL}}_{\mathrm{ef}}$	−1020.14	−3.97	−4.25	−48.89	−180.93	−9.02	−1.05	−4.02